Running Head: COMPOSITE FACE PROCESSING

Composite faces are not (necessarily) processed coactively: A test using Systems Factorial Technology and Logical-Rule Models

Xue Jun Cheng

Callum J. McCarthy

The University of Melbourne

Tony S. L. Wang

Brown University

Thomas J. Palmeri

Vanderbilt University

Daniel R. Little

The University of Melbourne

Address Correspondence to:

Xue Jun Cheng or Daniel R. Little

Psychological Sciences

The University of Melbourne

Parkville VIC 3010

xjcheng@student.unimelb.edu.au or daniel.little@unimelb.edu.au

http://www.psych.unimelb.edu.au/research/labs/knowlab/

Abstract

Upright faces are thought to be processed more holistically than inverted faces. In the widely-used composite face paradigm, holistic processing is inferred from interference in recognition performance from a to-be-ignored face half for upright and aligned faces compared to inverted or misaligned faces. We sought to characterize the nature of holistic processing in composite faces in computational terms. We use logical-rule models (Fifić, Little, & Nosofsky, 2010) and Systems Factorial Technology (Townsend & Nozawa, 1995) to examine whether composite faces are processed through pooling top and bottom face halves into a single processing channel – coactive processing – which has one common mechanistic definition of holistic processing. By specifically operationalising holistic processing as the pooling of features into a single decision process in our task, we are able to distinguish it from other processing models that may underlie composite face processing. For instance, a failure of selective attention might result even when top and bottom components of composite faces are processed in serial or in parallel without processing the entire face coactively. Our results show that performance is best explained by a mixture of serial and parallel processing architectures across all four upright and inverted, aligned and misaligned face conditions. The results indicate multi-channel, featural processing of composite faces in a manner inconsistent with the notion of coactivity.

Keywords: face processing, serial vs parallel, coactivity, categorization, Systems Factorial Technology

A hallmark of research on face processing is that upright faces seem to be processed in a qualitatively different manner compared to non-face objects (Scapinello & Yarmey, 1970; Yin, 1969). This conclusion has been extended to include differences in processing between upright faces and inverted faces and between aligned faces and misaligned faces (Chance & Goldstein, 1981; Rossion & Boremanse, 2008). For example, the famous Thatcher illusion (Thompson, 1980) shows that the same distortions in a face are difficult to detect in an inverted face but obvious in an upright face. In addition, recognition performance is less accurate and slower for inverted faces or misaligned faces than for upright faces (see e.g., Chance & Goldstein, 1981). A large number of compelling demonstrations of this sort have been identified, highlighting differences in face recognition depending on various forms of face presentation.

Holistic Processing

The concept of holistic processing has been proposed to explain the apparent difference between recognizing upright faces and recognizing other classes of common objects (including inverted or misaligned faces). It commonly refers to the notion that an object's features are simultaneously integrated into a single perceptual representation such that each object is represented as an emergent whole rather than by its parts, or indeed by the sum of its parts (Richler, Palmeri, & Gauthier, 2012; Rossion, 2008; Shen & Palmeri, 2015).

Two paradigms widely used to test for holistic processing are the *part-whole paradigm* (Farah, Tanaka, & Drain, 1995; Tanaka & Farah, 1993) and the *composite face paradigm* (Hayward, Crookes, Chu, Favelle, & Rhodes, 2016; Richler, Tanaka, Brown, & Gauthier, 2008; Rossion & Boremanse, 2008; Young, Hellawell, & Hay, 1987). The purpose of this paper is to characterize the recognition mechanisms used to process upright, inverted, aligned, and misaligned face stimuli in composite face tasks using methods that allow us to differentiate a large number of different processing strategies. We begin by introducing both the part whole and composite tasks, noting how these tasks identify holistic processing, and follow by introducing concepts that allow us to make more detailed temporal predictions that one should expect if a face is processed holistically.

Part-whole paradigm. The part-whole paradigm (Tanaka & Farah, 1993) was developed based on the notion that processing an entire face should show *configural superiority* over and above the summed information of each of its parts. Holistic processing in this task is defined as better recognition of the whole face rather than parts of the face, scrambled parts of the face, or the parts in a new context. In this task, participants are first trained to recognize members of different classes of stimuli (e.g., upright faces and scrambled faces) that differ on a single feature such as the eyes. Recognition of the target feature (in this case, the eyes) is then measured either in the context of the study stimulus (i.e., the specific face – e.g., Mary – that participants were trained on), in isolation, or embedded in a different context (e.g., Mary's eyes on Jane's face). Tanaka and Farah argued that it is possible to differentiate featural processing (independent processing of each part) from holistic processing (the face as a whole) by comparing recognition performance between conditions. Specifically, there should be a difference in recognition for a target feature in its studied upright face context compared to in isolation but no difference for scrambled faces compared to in isolation. Indeed, Tanaka and Farah found evidence of configural superiority for upright faces but not scrambled faces.

In a subsequent study, Farah, Tanaka, and Drain (1995) employed the part-whole paradigm to investigate differences in processing between upright and inverted faces. They reported a configural superiority effect for only the upright faces even though the same faces were used in both upright and inverted conditions. Farah et al. argued that inverting a face may disrupt second-order configural cues (i.e., relational spatial information between features, such as the distance between the mouth and the nose; Bartlett & Searcy, 1993) available in upright faces. As configural cues are important for holistic processing, due to their providing additional information to aid recognition (Schwaninger & Mast, 2005), inverted faces are not processed holistically in the part-whole task as they lack these cues.

Composite face paradigm. In the composite face paradigm (Young et al., 1987), novel composite faces are formed by aligning top and bottom halves from different people’s faces (for a review of different versions of the composite task, see Gauthier & Bukach, 2007). Young et al. showed that recognition of either face half when embedded in a composite is poorer than recognition of that half in its original configuration (called the *composite face effect*). This poorer identification is assumed to arise from an inability to attend to the top or bottom half of a face while ignoring the other half of a face (Richler et al., 2012; Rossion, 2008).

This composite face effect disappears when the halves of the composite faces are misaligned. Even though the misaligned faces have the same visual information as the aligned faces, spatially offsetting the top and bottom halves leads to better recognition of one half while ignoring the other half. The difficulty in selectively attending to one half of a face presented in its aligned configuration relative to a misaligned configuration has been interpreted as evidence for some form of holistic face processing (Cheung, Richler, Palmeri, & Gauthier, 2008; Farah, Wilson, Drain, & Tanaka, 1998).

In the original Young et al. study, participants were asked to recognize by name the top or bottom half of a composite face. In many contemporary classroom demonstrations of the paradigm, this might be asking someone to recognize the top half of a composite as Matt Damon's face while ignoring the bottom half of George Clooney's face. Most contemporary versions of the composite face task have used a same-different task whereby a study composite is shown briefly, and then, after a short delay, a test composite is shown and the participant must say where the top half (or the bottom half) of the test face matches that of the study face.

In one adaptation of the composite face paradigm, Rossion and Boremanse (2008) presented different bottom halves of faces with the same top half, with the two halves either aligned or misaligned. Recognition performance (judging half of the test face as same or different from the study face) was poorer in the aligned condition than in the misaligned condition. Performance in the aligned condition was also worse for upright faces compared to inverted faces, suggesting that aligned upright faces, but not misaligned or inverted faces, are processed holistically (but see Richler, Mack, Palmeri, & Gauthier, 2011).

More recent versions of the composite face task employ a slightly elaborated, *complete* design, in which changes to the top and bottom half between study and test can be either congruent or incongruent with one another (see e.g., Cheung et al., 2008; Gauthier & Bukach, 2007). In the *partial design* (as used by Rossion and Boremanse, 2008), the to-be-ignored half was always different regardless of whether the to-be-attended half was the same or different. Poorer performance in the partial design could consequently be due to a response bias (see Gauthier & Bukach, 2007). The complete design introduces both congruent changes (the to-be-ignored half is the same when the attended half is the same) and incongruent changes (the to-be-ignored half is different when the attended half is the same). If there is interference from the to-be-ignored half, then one predicts an interaction between congruency and alignment (or inversion, depending on the manipulation) in the complete design. Misalignment makes it easier to ignore the irrelevant half, consequently resulting in a reduction in sensitivity for incongruent changes because there is effectively only one piece of information rather than two – the top or bottom face half – signalling a change. For the aligned trials, there is both an effect of being able to use the irrelevant dimension on the congruent trials to boost performance and a failure to ignore the irrelevant dimension on the incongruent trials to harm performance. This interaction has been observed in a many experiments (e.g., Cheung et al., 2008; Chua, Richler, & Gauthier, 2014, 2015; Richler et al., 2011; Richler et al., 2015).

Limitations of part-whole and composite face paradigms. Most researchers interpret the reduction in recognition performance for upright aligned faces compared to scrambled, inverted, or misaligned faces as indicative of holistic processing (e.g., Hayward et al., 2016; Richler et al., 2008; Rossion & Boremanse, 2008; Young et al., 1987). One challenge is that the concept of holism has numerous meanings and the precise meaning of holism is not uniformly identifiable from the results of these tasks (Richler et al., 2012). Many different definitions of holism could explain the results:

participants could make the wrong decision about parts or halves of an upright faces due to unintended retrieval of a whole face template, disruption of the geometry of configural aspects of the face when combined in a different face context, or failure of selective attention. And notions of independence, say between face halves or face parts, are sometimes confused with notions of separability (as pointed out by Ashby & Townsend, 1986; Fitousi, 2015). Unlike standard empirical demonstrations of the part-whole and composite face paradigms, where holism is only inferred from reduced recognition performance, we adopt a double factorial paradigm that allows us to determine the precise nature of processing underlying those faces through the identification of information-processing architecture.

Our work follows on that of Fitousi (2015) and Fifić and Townsend (2010). Fitousi examined processing of composite faces comprised of top and bottom face halves using a composite face task, a series of unidimensional categorization tasks (following Garner, 1974) and a redundant target detection task. The composite face task was employed to ensure that the faces showed the necessary composite face effect. Having shown this effect, the Garner tasks were used to determine whether composite faces exhibited characteristics classically associated with stimulus integrality.

Stimulus integrality is a concept closely aligned with the idea of holistic processing. Integrality refers to stimulus dimensions that are difficult to attend to in isolation. Instead, performance with integral stimuli appears to be better described as processing of the whole object. In contrast, separability refers to dimensions that can be attended to easily in isolation. There are a number of converging operations demonstrating a difference between integral and separable dimensions (see Griffiths, Blunden, & Little, 2017, for a recent review) including, as described later, the applicability of different distance metrics to explain similarity data via multidimensional scaling (Euclidean for integral and city-block for separable; Nosofsky, 1992) and the utility of different processing architectures for explaining response times (namely, coactive processing for integrality and serial or parallel processing for separability; see e.g., Fifić, Nosofsky, & Townsend, 2008; Little, Nosofsky, Donkin, & Denton, 2013). Fitousi (2015) also applied another important test of integrality: Garner's selective attention tasks.

The classic Garner tasks are a set of selective attention tasks whereby a baseline condition is compared to two other conditions, a correlated condition and a filtering condition. In the baseline condition, two items that vary only on a single relevant dimension (with other dimensions held constant) are presented for the observer to categorize. In the correlated condition, there are again two items but these items have correlated variation on relevant and irrelevant dimensions. In the filtering condition, four items are presented with two items per category, with one dimension relevant for categorization and the other dimension irrelevant. We direct the reader to a recent comprehensive review by Algom and Fitousi (2016) and to our own recent work on this task (Little, Wang, & Nosofsky, 2016).

The Garner tasks identify integrality by determining whether categorization response times are (1) faster than baseline when there is correlated variation along the irrelevant dimension and (2) slower than baseline when there is irrelevant variation in the filtering task. Detection of facilitation and interference effects implies that the dimensions are integral, which is commensurate with certain notions of holistic processing (Pomerantz & Pristach, 1989). By contrast, with separable dimensions, irrelevant dimensional variation is ignored and performance is the same in all three conditions. In Fitousi (2015), although a composite face effect was observed, neither facilitation nor interference in the Garner task were observed. Richler, Palmeri, and Gauthier (2015) have similarly argued that the Garner tasks and the composite face task may tap different constructs of holistic processing and, therefore, it may be reasonable to expect different results. We return to this idea in our General Discussion.

Fitousi (2015) also used a redundant target task to assess workload capacity. He predicted that if the top and bottom face halves were processed holistically, then these composites should show a target detection benefit when presented together (compared to presented in isolation) over and above a mere statistical facilitation (the minimum time prediction) provided by a baseline independent race model (Raab, 1962). Observing that processing is faster than the minimum time prediction is sometimes termed supercapacity and is consistent with a notion of coactivity (Miller, 1982; Townsend & Eidels, 2011; Townsend & Nozawa, 1995). To explain, a coactive model predicts that redundant information is pooled into a common processing channel, providing one formal definition of the concept of holistic processing (see also Farah et al., 1998; Gold, Mundy, & Tjan, 2012; Maurer, LeGrand, & Mondloch, 2002; Tanaka & Farah, 1993; Young et al., 1987). Consequently, a coactive model predicts that redundant targets should be processed even faster than the baseline minimum time prediction of an independent race model. Although Fitousi observed a composite face effect, he did not observe supercapacity. In fact, capacity was limited.

In summary, when measured using tasks which assess independence (Garner) or processing capacity (redundant target detection), composite faces do not appear to show evidence consistent with the notion of coactivity, one important theoretical characterization of holistic processing. However, we note that because the Garner tasks are selective attention tasks, they do not allow for a direct measure of coactivity (see Little et al., 2013, for a discussion of this point). Instead, we require a divided attention task which allows us to factorially manipulate the difficulty of processing of both the top and bottom face halves (see Algom, Fitousi, & Eidels, 2017). Furthermore, the measure of capacity relies on the assumption of context invariance (Townsend & Nozawa, 1995). That is, the face halves must be processed at the same rate both in isolation and when presented together. This assumption may be readily violated in practice. Hence, in the present paper, we expand on these past approaches by moving beyond questions of independence to ask whether faces are processed according to a specific type of architecture using a measure that does not require context invariance to hold.

Certain architectures map to certain types of processing. For instance, in our previous work applying systems factorial methodologies to differentiate processing architecture (Townsend & Nozawa, 1995), we found that integral dimension stimuli, such as Munsell colors created by varying brightness and saturation, are processed coactively (Fifić et al., 2008; Little et al., 2013). Hence, coactivity seems to accord well with certain notion of holistic processing. By contrast, with separable dimension stimuli, such as shape and size (Moneer, Wang & Little, 2016), we found little evidence of coactivity and instead found processing consistent with serial or parallel processing (or, typically, a trial-by-trial mixture of serial and parallel processing; Fifić, Little, & Nosofsky, 2010; Little, Nosofsky, & Denton, 2011; Moneer et al., 2016). The notion of coactivity or pooling across the top and bottom face halves also seems to accord with the idea of holism being emergent from the sum of the different face parts, facial geometry, or retrieval of face templates. And the retrieval of face templates is consistent with the notion of exemplar-based processing (see e.g., Nosofsky & Palmeri, 1997). Fifić, Little, and Nosofsky (2010) showed that an exemplar-based model makes predictions that are formally identical to a coactive processing model.

An alternative definition of holism is provided by the idea that holistic processing represents a learned, strategic allocation of attention (Richler et al., 2012; Wong & Gauthier, 2010). The interaction in the complete composite face design, a hallmark of holistic processing, reflects effective use of attention or a failure to disengage attention selectively from specific face parts (Chua et al., 2014, 2015; Richler et al., 2011; Richler, Wong, & Gauthier, 2011). However, this failure of selective attention hypothesis can make predictions consistent with either coactive processing or serial or parallel processing.

To clarify these theoretical characterizations, we use the term processing architecture to refer to the organisation of mental processes (Kantowitz, 1974; Schweickert, 1992; Sternberg, 1969; Townsend, 1984). Serial processing indicates sequential processing of stimulus dimensions (the top half of the face is processed followed by the bottom half) and parallel processing indicates simultaneous processing of dimensions (both top and bottom are processed at the same time). These models can be further defined by their stopping rule. With a self-terminating stopping rule, processing stops whenever a decision can be made, even if only one dimension has been assessed. With an exhaustive stopping rule, processing continues until both dimensions have been assessed, even if a decision could have been made earlier. Both serial and parallel architectures are multi-channel processing models, where each dimension is processed independently. They can be contrasted with the coactive architecture, where both dimensions are pooled into a single channel that then drives a decision-making process[1]. Crucially, the mechanisms of the coactive architecture are analogous to the pooling aspect of certain characterizations of holistic processing: each stimulus feature is integrated into a single channel and processed together instead of being analysed independently.

Measuring processing architecture. Our methodology follows closely from Fifić & Townsend (2010) who examined processing architecture in the categorization of second-order features, such as the distance between the eyes or the distance between the nose and the mouth. We describe the details of Systems Factorial Technology (SFT) methodology in the following section, but to aid our discussion, we refer the reader to Figure 1. This figure shows the two conditions used by Fifić & Townsend (2010) as well as the condition used in the present experiments. In their OR task, a target category member can be identified by noting that either the eyes were closer *or* that the lips were higher than the sole contrast category member. In their AND task, a target category member could only be categorized if both features were different from the contrast category members features. Hence, responding in the AND task had to be exhaustive (in order to be accurate).

Fifić and Townsend tested interaction contrasts in both conditions when the faces were presented in the same context that they were learned, in a new face context, and in isolation (an *isolated features condition*, as in the part-whole paradigm). As described later, these interaction contrasts allow one to distinguish between serial and parallel, self-terminating and exhaustive models as well as coactive processing models.

[1] The coactive model can be thought of a parallel facilitatory model with perfect facilitation between the channels (see Eidels, Houpt, Altieri, Pei, & Townsend, 2011).

For their OR task, processing was most consistent with parallel, self-terminating processing in the old and new face contexts. In their isolated features OR condition, processing varied between parallel and serial self-terminating processing across observers. For their AND condition, however, some observers did show contrasts consistent with coactivity but only in the old face context; the remaining participants were more consistent with serial processing in the old face context and either serial or parallel processing in the remaining conditions. Nonetheless, Fifić and Townsend's experiment did provide some indication that, at least for some observers, there were consistent differences between part and whole conditions that were consistent with the notion that holistic processing reflects coactivity. The fact that coactivity was not invariant across tasks may implicate a strategic attentional mechanism (rather than more monolithic holistic pooling irrespective of task).

In the present experiments, we apply a similar methodology but to the composite face task by using factorially-manipulated top and bottom face halves as our set of stimuli. We test not only upright and aligned faces but also inverted and misaligned faces. Our methodology differs from Fifić and Townsend (2010) in two substantial ways: First, we extend our analysis to not only the target category items but also the contrast category items. Fifić et al. (2010; see also Cheng, Moneer, Christie, & Little, 2017; Little et al., 2015) showed that a large number of processing models make unique predictions regarding the contrast category items; these items add converging evidence as to the underlying processing architecture. Second, we complement the nonparametric interaction contrast SFT methodology with computational modelling. This allows us to test a number of novel explanations of processing beyond the canonical models which make up SFT. For example, we parametrically test mixtures of processing models (see e.g., Little et al., 2011). And we provide strong tests of coactivity by systematically increasing the flexibility of the coactive model (see e.g., Fifić et al., 2010).[2] In the following section, we outline how the use of interaction contrasts via SFT on target category stimuli and the contrast category allow us to differentiate processing models in our design (see Figures 2 and 3). We then present our experiment.

Systems Factorial Technology (SFT)

A critical feature of our design is that it links face categorization performance with Systems Factorial Technology (SFT; Fifić et al., 2010; Little, Altieri, Fifić, & Yang, 2017; Townsend & Nozawa, 1995; Townsend & Wenger, 2004, see Figure 1). There are several good recent tutorial reviews of SFT (Algom, Eidels, Hawkins, Jefferson, & Townsend, 2015; Altieri, Fifić, Little, & Yang, 2017; Harding et al., 2016). We provide a capsule summary of the relevant aspects of SFT we utilized.

The current study focuses on two sets of SFT analyses: the Mean Interaction Contrast (MIC) and Survivor Interaction Contrast (SIC). To explain how these interaction contrast analyses can differentiate processing architectures, note that the four members of the target category (x_1y_1, x_1y_2, x_2y_1, and x_2y_2) vary according to their discriminability (relative to the category boundary) on both dimensions. Stimuli closer to the category boundary should be harder to distinguish from the contrast category and are described as having low discriminability (L). Correspondingly, stimuli further from the boundaries should be easier to distinguish and are described as having high

[2] Our implementation of the coactive model assumes that there is no correlation in the perceptual representation of the top and bottom halves (i.e., perceptual independence; Ashby & Townsend, 1986). We test a highly-flexible model which allows a freely varying drift rate for each face. This model contains models which allow violations of perceptual independence as a special case.

discriminability (H). With regards to Figure 1, stimulus x_1y_1 is hereafter referred to as the LL stimulus, x_1y_2 as LH, x_2y_1 as HL, and x_2y_2 as HH. The interaction contrasts can be computed by combining these factorially-manipulated dimensions both at the mean and distributional level. Both interaction contrasts provide a powerful tool for differentiating processing models.

Mean Interaction Contrast (MIC). In our experiment, the target category faces require both dimensions to be processed, and hence are categorized using an AND rule. Distinct patterns of mean RTs for the target category items are illustrated in Figure 2 (left column). The mean RT interaction can be summarized as:

$$MIC = (RT_{LL} - RT_{LH}) - (RT_{HL} - RT_{HH}) \,. \qquad (1)$$

As shown in Figure 2, the serial, parallel, and coactive architectures each make a characteristically different prediction. Here we focus only on the exhaustive response patterns because self-termination in the target category would lead to a high error rate in our design.

The serial architecture predicts an additive pattern of mean RTs (MIC = 0). To explain, when both dimensions are processed exhaustively, the overall RT predicted by a serial model is given by the sum of processing both dimensions. Hence, the increase in RT from HH to HL should be approximately equal to the increase in RT from LH to LL. A parallel architecture predicts an under-additive pattern (MIC < 0). With an AND rule, the parallel architecture's predictions are given as the maximum processing time taken to process both dimensions. Hence, the processing time of the LH, HL, and LL stimuli will be much closer to each other than to the HH stimulus. For the coactive architecture, the pattern of RTs is predicted to be overadditive (MIC > 0; Fifić et al., 2010; Houpt & Townsend, 2011; Townsend & Nozawa, 1995). The intuition is that because information from both channels is pooled together, a single low discriminability value does not act to slow the entire RT; however, processing is slowed considerably when both dimensions have low discriminability.

Survivor Interaction Contrast (SIC). The survivor interaction contrast (SIC) allows us to gain further diagnostic information by analysing the entire RT distribution for each target category item rather than merely the means. The SIC is computed analogously to the MIC but uses the survivor function, S(t), for each item. The survivor function is defined as the probability that a randomly observed time *T* takes longer than *t* time units to complete: $S(t) = p(T > t)$; the survivor function is just 1 minus the cumulative distribution function often used to portray RT distributions. When $t = 0$, $S(t) = 1$ because all processing times exceed zero, and as t approaches infinity, S(t) moves toward zero. Slower processing is associated with greater values of the survivor function across the time domain. The SIC is given as:

$$SIC(t) = \left[S_{LL}(t) - S_{LH}(t)\right] - \left[S_{HL}(t) - S_{HH}(t)\right]. \quad (2)$$

Like the MIC, the SIC produce qualitatively distinct patterns for each of the processing architectures (see Figure 2, right column). Serial architectures are characterised by having equal positive and negative areas under the curve, parallel architectures by an entirely negative curve, and coactive architectures by a small negative deflection that then becomes largely positive. The integral of the SIC gives the value of the MIC (Townsend, 1990); hence, the MIC is a measure of the positive and negative area of the SIC.

Contract category predictions. The mean RTs from the contrast category items provide an extended set of qualitative contrasts for distinguishing between the architectures (Fifić et al., 2010). RT patterns from the contrast category items for all the architectures are necessarily exhaustive due to the logical rule defining their categorization (both dimensions must be processed to confirm categorization into the target category). Unlike the target category items, RT patterns from the contrast category items can reveal whether processing is governed by an exhaustive or self-terminating rules for the serial and parallel models. Detailed intuitive descriptions of the qualitative predictions and their rationale can be found in Fifić et al. (2010, pp. 313-317).

For instance, consider the difference between fixed-order and mixed-order serial self-termination processing. For a fixed-order serial self-terminating process that processes dimension x before dimension y, the mean RTs for the first processed dimension are approximately equivalent (see Figure 3, top left panels). For the second processed dimension, however, the mean RTs for the interior stimulus (e.g., x_1y_0, see Figure 1) are slower than the mean RTs for the exterior stimulus (e.g., x_2y_0); we use the terms interior, I , and exterior, E, to refer to a stimulus's position in the category space (for reference see Figure 1). The reason for this is because after processing dimension x, which satisfies the vertical rule bound in favour of the target category, processing must switch to dimension y. Because it is harder to discriminate x_1 from x_0 than it is to discriminate x_2 from x_0, this switch will occur later for the interior stimulus than the exterior stimulus. By contrast, for mixed order serial self-terminating processing (Figure 3, middle left panel), where either x or y may be processed first from trial to trial, both interior stimuli are predicted to have longer RTs than both exterior stimuli. This is because on some trials, depending on which dimension is processed first, processing will have to switch to the other dimension, and, again, this is will occur later for the interior stimuli than for the exterior stimuli.

A key contrast to these predictions is provided by the coactive architecture. Now, the interior stimuli are predicted to be processed faster than the exterior stimuli. The intuition is that in pooling across both dimensions, the interior stimuli pool more evidence for a contrast category response than the exterior stimuli since they are closer to the lower left-hand corner of the stimulus space (i.e., the contrast category). In sum, the pattern of RTs from both categories offer a rich set of additional diagnostic data that allow differentiation of (a) architecture, (b) stopping rule, and (c) for serial processing, the order in which the dimensions are processed.

Behavioral Experiment

The aim of our present experiment is to investigate the architecture underlying the processing of composite faces. In separate conditions, we focus on the processing of upright and inverted, aligned and misaligned faces. If composite faces are processed holistically when upright and aligned but not in any of the other conditions, we should find coactivity for upright aligned faces but independent (serial or parallel) processing for misaligned or inverted faces. This result would be consistent with some of the existing literature that suggests that inverted and misaligned faces are not processed holistically.

On the other hand, Richler et al. (2011) showed that inverted (aligned) faces in a composite face task also show the hallmarks of holistic processing in a composite face task. But for inverted faces, the interaction effect (in the complete composite design) takes longer (in time) to emerge. In the present experiment, where stimuli are displayed until a response is made, we might expect also to see coactivity in the inverted aligned condition. Consequently, our approach poses and tests a series of strong hypotheses about the notion of holistic processing. We summarize these predictions in Table 1.

General Method

Participants. Nineteen participants from the University of Melbourne community were randomly assigned to the upright aligned (N = 5), upright misaligned (N = 5), inverted aligned (N = 5), and inverted misaligned (N = 4) conditions. One participant from the upright aligned condition and one participant from the upright misaligned condition were removed from further analysis due to error rates over 30% for one of the items. Participants received $6 for each session plus a $2 bonus per session for accurate performance (≥ 90%). Testing humans was approved under Melbourne Human Research Ethics Committee 1034866.

Stimuli and apparatus. A 3 × 3 matrix of faces was created using a field morphing technique (Steyvers, 1999). The base face halves used in this experiment are shown in Figure 1 (middle panel) for the upright conditions (Goldstone & Steyvers, 2001, Experiment 1; Kayser, 1997). The inverted conditions were created by rotating these components 180 degrees. Top face halves were formed by morphing faces 1 and 2, and bottom face halves were formed by morphing faces 3 and 4. The top and bottom halves were then combined to create the nine face stimuli. Thus, each face in the stimulus space can be defined by a combination of values on the top and bottom face halves (e.g. the face in the top left corner of the space in Figure 1 is 100% Face 1, 0% Face 2, 0% Face 3, and 100% Face 4).[3]

Each composite face was presented upright or inverted in either an aligned or misaligned fashion (between conditions; see Figure 1). The stimuli were presented on a 1280 × 1024 resolution monitor set at a viewing distance of 70 cm. Each face subtended a visual angle of 4.66° × 2.46° in the aligned condition, and 4.66° × 3.93° in the misaligned condition. The gap between the face halves subtended a visual angle of 0.16° in all conditions. A Gaussian filter was applied to blur the edges around each of the faces. RTs were collected using a calibrated response box (Li, Liang, Kleiner, & Lu, 2010).

Procedure. Each participant completed eight one-hour sessions on consecutive or near-consecutive days over a period of two weeks. Participants were instructed to respond accurately but also told that RTs were being recorded. At the start of each trial, a fixation cross was presented for 1500 ms. A single face was then presented and the participant was required to indicate whether the face belonged to the Target Category or the Contrast Category by pressing the appropriate button on the response box. Feedback was provided only if the response was incorrect (i.e., "…WRONG…") or after 5000 ms when the trial timed out (i.e. "…TOO SLOW…"). A blank interval of 2000 ms was inserted between trials.

[3] In our initial pilot tests, the upright condition was much easier than the inverted condition such that there were only minimal RT differences between the L and H levels. This precludes the analysis of the SFT contrasts since we did not find effective influence of the manipulations of the top and bottom face halves (Townsend & Nozawa, 1995). To increase the difficulty of the upright condition, rather than use 0%, 50%, and 100% morph levels on each dimension, we used 20%, 40% and 80%.

Each face was presented 50 times per session for a total of 400 times per participant across all eight sessions. In each session, there were ten blocks of 45 trials each, with the presentation order of the faces randomised within each block. Between blocks, participants were allowed to take a short break and given feedback on their performance accuracy for the previous block.

Data Analysis. Following related precedents using SFT, we focus our analysis on individual participant ANOVAs to assess the target category results and on planned *t* tests to assess the contrast category results. These analyses allow us to make inferences about the specific processing patterns for each individual participant. One shortcoming of this method is that because our manipulation of face alignment and direction is between subjects we provide no direct comparison between conditions. However, our focus is on the specific processing architectures arising for each participant in each condition. Our non-parametric and parametric modelling analyses allow for a consistent interpretation across conditions.

Results

In all conditions, session one was considered practice and excluded from analysis. For each participant and each item, trials with RTs less than 200 ms or greater than three standard deviations above the mean were excluded. This resulted in less than 2% of trials being removed. The mean correct RTs, mean error RTs, and error rates for each participant are presented in Tables 2 and 3 for the upright and inverted conditions, respectively. As shown, error rates tended to be low across all the items for most participants with the exception of the LL stimulus and the Ix and Ex stimuli. We refer to participants by an abbreviation of their assigned condition (e.g., UA1 is the first participant in the Upright Aligned condition).

Before turning to the individual conditions, we note some results which were consistent for all participants in all conditions. To analyse the target category, we conducted a 7 (session: 2-8) × 2 (top half: L or H) × 2 (bottom half: L or H) ANOVA on the target-category RTs for each individual participant (see Table 4 and Table 6). These results show that: (1) the main effects of top- and bottom-half discriminability were significant for all participants in all conditions indicating that our manipulation was successful. (2) There was a main effect of session for all participants indicating that RTs became faster with practice. (3) For some subjects, session interacted with one or both dimensions indicating that one or both dimensions became processed more quickly relative to the other dimension across trials. (4) The three-way interaction between session, top, and bottom half was not significant for any observers, indicating that the relationship between the target category items was stable across sessions; the sole exception to this was participant UM4. The three-way interaction indicates that the relationship between the target category items was unstable across sessions. A possibility is simply that performance had not stabilized by the second session. Since we wish to examine only asymptotic behavior, we examined the data after removing session 2 (i.e., in addition to session 1), in which case the three-way interaction was no longer significant. These are the analyses reported in Table 4.

For the contrast category, following Fifić et al. (2010), we conducted a series of planned *t* tests comparing the redundant stimulus to each of the other contrast category stimuli and comparing the interior to the exterior stimuli on both dimensions. These comparisons showed that for all of the subjects, the redundant stimulus was faster than the remaining stimuli ruling out any exhaustive processing model. For all participants, with the exception of UA4 and UM2, the top half of the face is processed faster than the bottom half of the face, regardless of whether the face is presented at the top or bottom of the display (e.g., in the upright and inverted conditions, respectively; see Tables 5 and 7).

Having covered the results that are consistent across all the conditions, we now turn to the two primary diagnostic measures: the two-way interaction of target and direction of that interaction (the MIC) between the top and bottom face halves for the target category stimuli and the interior versus exterior comparison for the contrast category stimuli. The pattern of results across these two measures allows us to rule out different model architectures describing how the top and bottom halves are integrated when making a category decision. To remind the reader, coactive processing would be identified by a significant, positive MIC, along with a mostly positive SIC, coupled with faster mean RTs for the interior than exterior contrast category items. Serial processing would be identified if the target category MIC is near zero (and non-significant) coupled with a slower interior than exterior contrast category items. Parallel processing would be consistent with a significant negative MIC, and a negative SIC, coupled with a non-significant interior/exterior comparison indicating no difference in the mean RTs.

Upright Aligned Condition

Target category. Figure 4 (top row) shows the SIC and mean RTs for the target category items for the upright aligned condition. Referring to the model predictions in Figure 2, participants in the upright aligned condition were most consistent with the underadditive predictions of the parallel model (or the serial model for participant

UA4). For participant UA3, the significant top x bottom interaction confirms that the MIC was negative supporting an inference of parallel processing (see Table 4). This interaction was not significant for the participants UA1, UA2, or UA4 which is suggestive of serial processing.

Contrast category. The contrast category analyses suggest some differences between participants (see Table 5 and Figure 5, top row). For UA1, the analysis indicates parallel self-terminating processing since there is no difference between the interior and exterior stimuli on either dimension consistent with parallel self-terminating processing. For participant UA2, there is no difference between the top face half, but the exterior is faster than the interior on the top face half. For participants UA3 and UA4, the exterior stimulus is faster than the interior stimulus for only the top dimension. For the bottom dimension, however, the interior stimulus tends to be faster than for the exterior stimulus, and the comparison is only significant for participant UA4. The contrast-category results for participants UA3 and UA4 are inconsistent with any of the models.

Summary. Taken together, for participants UA1 & UA2, the SIC appears to be underadditive, but the interaction contrast is non-significant. When coupled with the contrast category results, the data are suggestive of parallel self-termination for participant UA1 but of mixed order serial-self termination for participant UA2. In previous work, this result has been commonly found when processing is best explained as a trial-by-trial mixture of serial and parallel processing (cf. Little et al., 2011; Moneer et al., 2016). For participant UA3, the SIC is mostly negative, and the MIC is significantly negative suggesting parallel processing. Participant UA4's target category results look serial, and this inference is supported by a non-significant MIC. However, for both of these participants, the top dimension shows longer exterior than interior RTs but the bottom dimension shows longer interior than exterior RTs.

On the surface, our contrast category results appear only partially consistent with coactivity. However, it may be possible to explain this by assuming that the perceived variabilities of the top and bottom halves are not equivalent across all values of each dimension. The implication is that due to interacting attentional demands to the top and bottom half and to items close to the dimensional boundary, some face halves may not be as well learned as others. This leads to increased perceptual variability for the exterior "bottom" item compared to the interior "bottom" item and would allow for the longer exterior RT. Though post hoc, this explanation is supported by the observations that (1) the top half is processed faster than the bottom in almost all cases, suggesting decreased variability for the top half and (2) the psychological representation of the upright condition shows that the items near the boundary are more difficult to discriminate due to the spacing of these items (see Appendix A). We explore this explanation further by fitting a flexible coactive model which estimates a freely varying rate of processing for each stimulus. If processing is truly coactive, then this model should allow us to account for these complex effects. We defer further discussion of this explanation to our Computational Modelling section. In sum, the nonparametric analyses reveal little evidence for coactivity in the upright aligned face condition.

Upright Misaligned Condition

Target category. For participants UM1, UM3, and UM4, the SIC results are most consistent with parallel self-terminating processing (see Figure 4). A negative MIC was observed for each participant, indicating parallel processing. To confirm, the top x bottom interaction was significantly negative for participant UM1 and UM4 and marginal for UM3 supporting this inference (see Table 4, right hand columns). This interaction was not significant for UM2 which is suggestive of serial processing.

Contrast category. There was no significant difference between the interior and exterior items for UM1 or UM3 adding further support to the inference of parallel self-terminating processing. For UM4, the interior is processed slower than the exterior stimulus for the top face half but this is reversed for the bottom face half.

For UM2, who showed target results consistent with serial processing, we see a similar patter to UA2 in which there is no difference between I and E for the slower of the two dimensions (for UM2, the top), but there is a difference on the faster of the two dimensions (the bottom) with the exterior being faster than the interior.

Summary. Taken together, the target and contrast category results provide clear evidence for parallel self-terminating processing for UM1and UM3. UM2's non-significant MIC, S-shaped SIC, and significantly longer interior than exterior RT on the bottom dimension suggest serial processing. The data for UM4 are more equivocal. The target category appears to be clearly parallel, but the contrast category shows some signs of faster interior than exterior processing as found for two of the participants in the upright condition.

Inverted Aligned Condition

Target category. Four of the five participants in the inverted aligned conditions show the positivity (MIC > 0) associated with coactive processing (see Figure 6). However, the top x bottom interaction is not significant for participants IA1, IA3, and IA4 (see Table 6). By contrast, IA5's significant positive interaction indicates an overadditive MIC which is consistent with coactivity. IA2's negative SIC and significant negative interaction indicates parallel processing.

Contrast category. Of the four participants whose target category results suggest coactive processing, only comparisons for IA1 show significantly shorter interior than exterior RTs on both stimulus dimensions (see Table 7 and Figure 7). IA5 show RTs consistent with coactivity but the interior is faster than the exterior only on the faces which satisfy the contrast category rule for the bottom dimension. For the remaining participants, the RTs on the interior and exterior stimuli are equivalent (see Figure 7) or the exterior stimulus RTs which are faster than the interior stimulus RTs (see e.g., IA4). This result supports an inference of parallel self-terminating processing for IA2 and IA3. Participant IA4's results indicate shorter RTs for the exterior top stimulus compared to the interior top stimulus. This could indicate serial self-terminating processing similar to Upright Aligned participant UA2.

Summary. Taken together, we find clear evidence for parallel self-termination (IA2) and, surprisingly, clear evidence for coactivity (IA1 and IA5; see Figures 6 and 7). We are unable to unambiguously determine processing for the remaining two participants based on the non-parametric results; however, as described below, the parametric modelling allows for a more conclusive decision.

Inverted Misaligned Condition

Target category. For all participants, the observed SIC results have a substantial positive component. For IM1 and IM2, these SICs are most consistent with coactivity (see Figure 6; bottom row panels). For IM3 and IM4, the SIC is also consistent with coactivity, but the increased area in the early negative component could also reflect serial processing. However, the two-way interaction was not significant for any participant and this implies that the MICs, though positive, were not differentiable from zero (see Table 6).

Contrast category. For the contrast category mean RTs, only IM2 showed a shorter RT for the interior than the exterior stimulus, and only for the bottom dimension (see Figure 7). Participants IM1 and IM4 had equivalent RTs on the interior and exterior stimuli for both dimensions consistent with parallel self-terminating processing (see Table 7). Participant IM3 had equivalent RTs on the redundant, interior, and exterior stimuli on the top dimension and longer interior than exterior RTs on the bottom dimension. These results are consistent with serial self-terminating processing.

Summary. Here again, the target and contrast categories offer a mixed set of results that do not allow for easy interpretation. Only IM3 can be classified on the basis of the non-parametric results. This participant shows clear evidence of serial and self-terminating processing. The remaining participants show SIC results that suggest coactivity but the observation is not confirmed by the statistical analysis. The contrast category results are most consistent with parallel self-termination. However, the results from each category are somewhat difficult to reconcile. We turn to computational modelling later to provide an avenue for examining models which violate some of the assumptions inherent in the non-parametric analysis.

Summary of SFT results

A challenge of SFT is that heterogeneity between subjects often precludes drawing general conclusions about processing architecture; however, we can draw on other aspects of the data to offer some interim explanations.

Most clearly, the non-parametric analyses of SFT do not provide strong evidence of coactivity in the upright aligned condition or any of the other conditions. Although there were a small number of subjects who showed faster interior than exterior stimulus mean RTs, this was never accompanied by an overadditive MIC or SIC in the target category. Oddly, the only evidence for any overadditivity in the target category was found in the inverted conditions.

We also found no evidence for exhaustive processing. That is, the redundant stimulus in the contrast category is always processed as fast as or faster than the interior or exterior stimuli (clearly, processing in the target category is forced to be exhaustive by the experimental design).

Most of the participants show results that resemble the predictions of the parallel or serial processing architectures (with some participants exhibiting predictions which seem to fall in between these two architectures' predictions).

To further clarify our results, we use parametric computational models. These allow us to test both mixtures of serial and parallel processing along with a more flexible model of coactivity. Although the error RTs are not utilized by the SFT analyses, we fit the error RTs using parametric computational models.

Computational Modelling

We fit a set of logical rule-based models that were introduced by Fifić et al. (2010) to differentiate potential model architectures on their ability to explain our observed data. These models have been applied to examine processing of a number of different stimulus dimensions during perceptual categorization (Blunden, Wang, Griffiths, & Little, 2015; Little et al., 2011; Little et al., 2013; Moneer et al., 2016). The logical rule models take their name from the fact that the outcome of independently processing multiple stimulus dimensions may combine using logical OR or AND operations. Each of the independent processes is modelled as an evidence accumulation process (Brown & Heathcote, 2008; Luce, 1986; Nosofsky & Palmeri, 1997; Ratcliff, 1978), whose predicted response times are then combine depending on the specific mental architecture model. For example, a serial self-terminating model would sum the two independent processing times, thereby following an AND rule. A parallel exhaustive model would use the maximum of the two independent processes, implementing an AND rule. The corresponding self-terminating processes would implement an OR rule for certain stimuli (such as the contrast category stimuli). Our nonparametric SFT analyses rule out exhaustive processing for the contrast category results. Consequently, we focus our analyses on the self-terminating models.

To briefly summarize, following General Recognition Theory (GRT; Ashby & Townsend, 1986) and the related decision boundary theory (Ashby & Gott, 1988), the logical rule models assume that the perception of any stimulus dimension (e.g., top or bottom face half) is noisy and repeated presentations of the same dimensional value may not result in the same percept (see Figure 8; panel A; see also Richler, Gauthier, Wenger, & Palmeri, 2008). We represent the distribution of percepts for each stimulus dimension as a normal distribution with a mean and variance. For serial and parallel models, samples are taken from the marginal distribution on each dimension and used to drive independent sequential sampling decision process (shaded distributions in panel A, Figure 8). So, for example, if the participant samples from the top face half and that face half is identified as belonging to the target category, then evidence is accumulated toward the target category.

For the coactive model, we assumed that each stimulus location is identified by a bivariate normal distribution of perceptual effects and that these perceptual effects, along the two dimensions (top and bottom face halves), are statistically independent for each stimulus.[4] For all models, it was also assumed that all stimuli have the same perceptual variability along the two dimensions but that this variability differed between dimensions. Consequently, separate free parameters, σ_{Top} and σ_{Bot}, were allowed for the standard deviation of the normal distribution for the top and bottom face-half

[4] Note that within the GRT framework, there are other possible definitions of holistic processing. For instance, a violation of perceptual independence or perceptual separability (i.e., mean shift integrality; Ashby & Townsend, 1986) provide alternative senses of holism. Since we assume independent channel serial and parallel models are driven by the marginal perceptual distributions, violations of perceptual independence would only matter for the coactive model. Since we use GRT as a "front-end" to derive a drift rate, our free-drift model allows for a test of violations of perceptual independence and separability in a coactive architecture (see e.g., Fific et al., 2010). We note, however, that there may be models with increased generality (e.g., Townsend, Houpt & Silbert, 2012; see Griffiths, Blunden & Little, 2017 for a discussion of this point).

dimensions, respectively. The means, μ, of the perceptual distributions were fixed to the coordinate values obtained from the multidimensional scaling (MDS) solutions for these specific face stimuli (see Appendix A). This analysis revealed that the pairwise similarity ratings for faces in each condition were best captured by a constrained grid-model using a city-block metric to capture the distance between stimuli. The need for a city-block distance metric itself is evidence that our composite faces are inconsistent with some notions of integrality (Melara, Marks, & Lesko, 1992; Nosofsky, 1992).

The models assume that the location of the decision-bounds on the top and bottom face-half dimensions, D_{Top} and D_{Bot}, respectively, are established by each participant in order to differentiate the target category from the contrast category. The evidence for each category is found by integrating the bivariate normal distribution within each of the category regions. This evidence is used as the rate of evidence accumulation, or drift rate, v, in the sequential sampling model.

Previous applications of the logical rule models have used a discrete time sequential sampling model, the random-walk model; the present application instead uses a simplified continuous model of RT, the Linear Ballistic Accumulator (LBA; Brown & Heathcote, 2008). Practically speaking, using the LBA instead of random walk or diffusion processes circumvents certain problems with optimising discrete and continuous random walk parameters. The various methods for modelling RTs are conceptually quite similar, and the LBA was chosen primarily for its greater tractability.

The LBA generates RTs by assuming that two accumulators race against each other with the faster of the two accumulators generating the RT (see Figure 8, panel B). A single accumulator’s finishing time is determined by the relationship between the distance between the starting point of accumulation and the threshold for making a decision and the rate of evidence accumulation is given by the relationship: time = distance/rate (a standard velocity function). To achieve variability in the RT distribution, the LBA assumes that the start point for each accumulator varies from trial to trial according to a uniform distribution (parameter A provides the range of this distribution) and that the drift rate varies from trial to trial according to a normal distribution (parameter *s* provides the drift rate standard deviation). LBA also has two threshold parameters, one for the target category accumulator, and one for the contrast category accumulator. Each model also assumes a non-decision time parameter, t_0, that captures time not associated with decision-making processes (processes such as encoding or motor-execution). The final predicted RT is the decision time predicted by the LBA plus the non-decision time.

For the serial and parallel architectures, there are separate accumulators for the top and bottom face halves (see Figure 8; panel C). By contrast, for the coactive model, there is a single accumulator for the entire face. Each pair of accumulators contains one accumulator corresponding to the target category and one accumulator corresponding to the contrast category. The drift rates for each pair of accumulators are determined by the marginal perceptual distributions along each dimension (see Figure 8, panel A). To determine the drift rates (e.g., the top face half dimension), we integrate the marginal distribution with respect to the decision boundary to determine the area under the curve that falls into each category region along that dimension; hence, the drift rates for the correct and error accumulators sum to one.

In total, there are nine free parameters across the models: the perceptual standard deviation parameters (σ_{Top} and σ_{Bot}); location of the decision-bounds on the top and bottom face-half dimensions (D_{Top} and D_{Bottom}); the range of the start-points for evidence accumulation (A); the correct and error thresholds (b_{Target} and b_{Bottom}); drift rate variability (s); and non-decision time (t_0). The serial self-terminating model requires an additional parameter, p_x representing the probability that, on a given trial, the top half is processed before the bottom half.

We also fit an additional independent channel model that assumed that for any stimulus, the top and bottom half may be processed in parallel on some of the trials and serially on other trials. This mixture of processing might represent fluctuations in controlled attention across the course an experiment (e.g., Schneider & Shiffrin, 1977). The mixed serial-parallel model also includes the p_x parameter (as in the serial self-terminating model). In addition, the mixed serial-parallel model includes another extra parameter, p_{Serial}, which represents the proportion of serial trials. To allow for differences between the parallel and serial components, we multiplied the perceptual variability parameters in the parallel component by a freely estimated constant, $\sigma_{Parallel} = m\sigma_{Serial}$. We also estimated a separate starting point parameter, $A_{Parallel}$, for the parallel model component.

To model coactivity, we assume that the stimulus is represented by a joint bivariate perceptual distribution. For this model, a single drift rate is estimated by integrating the bivariate distribution with respect to both decision boundaries. To provide a more powerful test of coactivity that was unconstrained by our assumptions of the locations or variances of each stimulus, we fit a *free-stimulus-drift* model which simply estimates, as a free parameter, the drift rate associated with each stimulus. If we find that our constrained independent channel models fit better than this highly flexible model, then this would provide a strong result arguing against coactivity.

Fitting Procedure

We implemented the logical rules models in a hierarchical Bayesian framework. To compute the likelihood for each item, we simulated 50,000 data points from the model and used Holmes's (2015) probability density approximation method (PDA; see also, Turner & Sederberg, 2014) to determine the likelihood of each choice and each observed RT. The likelihood of the parallel, and coactive logical rule models (LR) is given the product over RTs and response on each trial:

$$L\left(D_{Top}, D_{Bot}, \sigma^2_{Top}, \sigma^2_{Bot}, A, b_A, b_b, s, t_0 \mid resp, rt\right) = \prod_{i=1}^{N} LR\left(resp_i, rt_i \mid D_{Top}, D_{Bot}, \sigma^2_{Top}, \sigma^2_{Bot}, A, b_A, b_b, s, t_0\right) \quad (1)$$

The serial self-terminating model has an addition parameter, p_x. The mixed serial-parallel model adds four additional parameters (p_x , p_{Serial}, m, $A_{parallel}$) and the free drift model replaces the GRT parameters ($D_{Top}, D_{Bot}, \sigma^2_{Top}, \sigma^2_{Bot}$) with freely estimated drift rates for each item ($v_{HH}, v_{HL}, v_{LH}, v_{LL}, v_{Ex}, v_{Ix}, v_{Ey}, v_{Iy}, v_R$).

We used Differential Evolution Markov chain Monte Carlo (DE-MCMC; Turner, Sederberg, Brown, & Steyvers, 2013) to efficiently generate proposals from the posterior distributions of each parameter. The variability in the likelihood approximation can result in "stuck" chains if an accepted parameter set results in an unusually high likelihood. To overcome this, we resampled the likelihood of any existing chains anytime the current proposal was rejected (see Holmes, 2015, p. 17). This resulted in good mixing between the chains and good convergence after an initial burn-in period. We used a burn-in period of 1150 iterations with a deterministic migration step (Turner et al., 2013) occurring every 20 iterations between iterations 501 to 700. For the remainder of sampling, we used a migration step instead of a crossover step with a probability of .05. For a handful of model fits, we increased the burn-in by taking an additional 1000 iterations. We used three times the number of parameters for each model to determine the number of chains. To improve sampling, parameters were transformed to lie in the range –∞ to +∞. A total of 750 iterations per chain were estimated giving, at a minimum, 20000 posterior samples per parameter. Prior distributions were selected to be relatively informative for each parameter placing each parameter in the range determined by previous applications of these models; the prior distribution settings are shown in Appendix B.

We used the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002) to compare each of the models. The deviance of a posterior sample of parameters, θ, is computed as $D(\theta) = -2\ln L(y\,|\,\theta)$. The DIC is computed as:

$$DIC = \bar{D}(\theta) + 2p_D \qquad (2)$$

where $\bar{D}(\theta)$ is the mean of the distribution of posterior deviances and $p_D = \bar{D}(\theta) - D(\bar{\theta})$. $\bar{\theta}$ is the average posterior parameter values. Thus, DIC penalizes the average negative log likelihood by a term which accounts for the functional form complexity of model.[5] The model that yields the smallest DIC is preferred. The DIC for each subject and each model is presented in Table 8.

---

[5] An alternative version of DIC uses $p_D = 2\,\mathrm{var}(\ln p(y\,|\,0))$ (Gelman, Hwang, & Vehtari, 2014). We tested this version as well; the results were the same. We also conducted a series of parameter and model recovery tests to compare both versions of DIC to the Watanabe-Akaike Information Criterion (WAIC; Watanabe, 2010) to determine how well each could recover the data generating model. Model/parameter recovery is useful because it allows the researcher to determine whether a given method is biased in its results (Heathcote, Brown, & Wagenmakers, 2015). Our results indicated that both versions of DIC recovered the generating model accurately, but WAIC was biased. In many simulations, the coactive model was selected as the preferred model by WAIC even though the simulated data were generated by the serial, parallel, or mixed serial-parallel model. We opted to use DIC for our model comparison analysis.

Summary of Model Fits

With few exceptions, the most preferred model was the mixed serial-parallel model. This model was preferred for two participants in the upright aligned conditions, all four participants in the upright misaligned condition, four of five participants in the inverted aligned condition, and three of four participants in the inverted misaligned condition. One participant in the upright aligned condition (UA4) was better fit by the serial self-terminating model. Hence, overall, we found little evidence of a difference between conditions.

Overall, we did not find strong evidence for coactivity in the upright aligned condition (or any condition for that matter). There were two subjects (one in the upright aligned condition, UA3, and one in the inverted aligned condition, IA5) who were best fit by the free stimulus drift model. The better fit of this model could indicate pooling across the top and bottom face halves for these participants.

We plot the 25, 50, and 75 percentiles for the correct and error RTs for each item and each subject in Figure 9. The models fit the data well and the parameter estimates were sensible. Detailed posterior predictive distributions for each subject and each item along with posterior parameter distributions are available in the supplementary materials.

Discussion

A summary of both the nonparametric and parametric results is provided in Table 9. While there are occasionally discrepancies that arise between the nonparametric and parametric results, we note that the parametric results consider both the correct and error RTs for all items simultaneously, unlike the nonparametric result. Consequently, we weight our discussion toward the outcome of our computational model fitting. Across all four conditions, our modelling results revealed the processing was largely consistent with a mixture of serial and parallel processing for a number of subjects. This type of mixture model has been found to provide good fits for a number of different stimulus types including overlapped separable dimensions (Little et al., 2011; Experiment 2) and whole object features (e.g., size, color, and saturation varied within a single object; Moneer et al., 2016). Hence, for many subjects, it appears they treated the composite faces no different from other objects with separable dimensions. There was some evidence of pooling (indexed by the free drift model) for two of the observers, but this appeared in both the upright aligned condition and the inverted misaligned condition.

The key point, however, is that there is not unequivocal evidence for coactivity (one standard definition of holistic processing) when participants categorize upright aligned faces. Nor is there clear evidence that upright faces are treated any differently from misaligned or inverted faces. Overall, there is no clean distinction between the processing of upright aligned faces and other faces. Any differences that we observed between upright and inverted, aligned and misaligned faces were largely quantitative rather than qualitative (see also Richler et al. 2011).

To ensure that face stimuli we used in our experiments yielded typical holistic processing effects, we conducted a standard composite face task using these stimuli and report the results in Appendix C. Our composite task results confirm that our upright faces (and our inverted faces) demonstrate an interaction between congruency and alignment, which is the hallmark of holistic processing in the composite face task (cf.

Richler et al., 2011). Although the faces we use appear to be processed holistically in the standard composite face task, we find no clear evidence that this is linked to coactivity in the categorization task (assess with either SFT or computational modeling).

General Discussion

We examined processing architectures underlying face processing using a set of composite face morphs in a categorization task. One operationalization of holistic processing is coactivity or pooling across the top and bottom face halves (Farah et al., 1998; Gold et al., 2012; Maurer et al., 2002; Tanaka & Farah, 1993; Young et al., 1987). Using a set of strong non-parametric and parametric inference techniques comparing a number of different processing models, we found little evidence for coactivity in the processing of upright aligned composite faces. Instead, processing of faces, whether upright or inverted, aligned or misaligned, consisted of mostly a mixture of serial and parallel processing.

Implications for theories of face processing

The interaction between congruency and alignment in the complete composite face design is one of the strongest pieces of evidence that upright aligned faces are processed holistically. However, as argued by Richler et al. (2012), the concept of “holism” can mean a variety of different things in terms of representations and processes. For example, holism can mean that upright aligned face have emergent properties beyond a simple summing of individual features, that upright aligned faces prompt the retrieval of holistic facial templates, that upright aligned face allows for use of stable configurations of features and their metric relationships, or that upright aligned faces allow for the use of an overlearned pattern of selective attention.

In the present article, we operationalize holism as coactivity, which implies a pooling of the top and bottom face halves into a single representation. This notion of coactivity acts as a proxy that accords well with several of the notions of holistic processing listed above, especially ideas of face templates and face configurality (see e.g., Fifić et al., 2010; Fifić & Townsend, 2010). Our findings of a lack of coactivity, and instead consistent serial/parallel processing across all conditions seems, we argue, more consistent with the ideas underlying the selective attention account of the composite face effect (Richler et al., 2012).

The selective attention account of the composite face effect argues that the interaction in the composite face task arises due to an inability to disengage from an overlearned pattern of attention when presented with an upright aligned face. While this theory does not provide strong constraints on processing, we feel that it is at least consistent with a parallel processing account of composite face performance. For instance, in the composite face task, one is instructed to ignore one half of the face. Misalignment improves one's ability to selectively attend only to the instructed face half. Congruent changes (where the top and bottom halves are either both the same or both different from a to-be-remembered face) result in improvement in sensitivity compared to incongruent changes. If the top and bottom halves are processed in parallel, then the effect of congruency likely occurs at the decision stage of processing. The implication is that there may be some confusion in the decision about whether a particular sample of evidence arose from the top or bottom face half. Our findings demonstrates that this confusion does not necessarily arise early in processing, which would have been observed as coactive processing.

Fifić and Townsend (2010) argued that there may be weaker definitions of holistic processing that can explain experiment effects typically used to infer holistic processing. For instance, there is a sense in which exhaustive processing may represent a form of weak holism. However, we found no evidence for exhaustive processing in our task (excluding, of course, the target category, where exhaustive processing is mandated by design). In some sense, parallel processing may also imply a weak form of holism since it can be used to explain tasks from other paradigms which demonstrate findings thought to be due to holistic processing.

For example, consider recent work using SFT to examine the other race effect. The other race effect occurs when recognition of faces of one’s own race is faster or more accurate than faces of another race. This is usually taken to indicate holistic processing of own race faces. Yang, Fifić, Chang, and Little (2017) demonstrated using a similar task to the categorization task used here that processing of own race faces was better characterized by a parallel self-terminating processing model whereas the processing of other race faces was better characterized by a serial self-terminating processing model. The increased efficiency afforded by parallel processing was sufficient to explain the other race effect as observed in the standard other race paradigm. That study also found no evidence of coactive processing.

The fact that we did not find evidence of coactive processing in our task does not rule out other dynamic models that involve pooling at an earlier stage of processing. For instance, initial encoding of the face may proceed coactively only for attention processes to break down the face into components at a later stage.[6] By manipulating divided attention and discriminability, we are necessarily concerned with whether processing is pooled or not during the later decision stage of processing. Examining pooling during encoding would likely require a different task with a manipulation of difficulty tailored to that particular stage of processing.

Relation to Previous Research on Composite Face Processing

As in the present task, Fitousi (2015) found a composite face interaction effect but no evidence of supercapacity (an indication of coactivity) and no evidence for Garner interference (an indication of integrality). Similar task-specific differences have been observed by Richler et al. (2015), which prompted those researchers to argue that the composite face task taps a different type of holistic processing from that tapped in phenomena such as Garner interference. By implication, this would also extend to Fitousi's redundant target task.

Richler et al.'s (2015) argument was based on a comparison of an interaction in the composite face task (the hallmark of holistic processing in that task) for faces differing in features but having the same configuration. This result contrasted with the earlier Amishav and Kimchi's (2010) finding that these same configuration stimuli did not show interference in a Garner filtration task, a result that they attributed to the independence of features and configurations. Richler et al. (2015) argued that the composite task is different from Garner interference for a number of reasons including generality (the composite task tends to be face specific) and susceptibility to stimulus

[6] We thank an anonymous reviewer for this suggestion.

discriminability (e.g., Melara & Mounts, 1993). As noted by Richler et al., even if both tasks measure failure of selective attention, it is not the case that either task can reveal anything about the source of that failure. The present task offers additional insight into the specific processing differences that occur in different tasks. By utilizing SFT, we are able to point quite clearly to the precise nature of processing that underlies performance on our task in processing composite faces. However, we find little evidence of (a) consistent holistic processing for upright aligned faces or (b) a real difference between upright and inverted faces.

Why did we not find coactivity for upright faces? Blaha (2017) recently reviewed a series of capacity studies showing that the evidence for the related concept of supercapacity for faces is mixed even within a single task. For instance, Ingvalson and Wenger (2005) found evidence for supercapacity with faces in a redundant target task, but Perry, Blaha, and Townsend (2008) and Donnelly, Cornes, and Menneer (2012) did not. Further, Wenger and Townsend (2006) found some evidence for supercapacity but only in some face conditions (specifically, faces shown in the studied context in the studied configuration). Blaha's careful critique identified a number of factors that may have contributed to these different results. In the present case, this analysis is informative because our categorization task is closely related to the redundant target task used in those experiments (Townsend & Nozawa, 1995).

One factor that seemed to contribute to a failure to find supercapacity is requiring subjects to make fine discriminations in order to differentiate faces. First, these discriminations might invoke selective attention, thus breaking the general facial gestalt. Second, certain tasks might emphasise reliance on the facial gestalt, thereby allowing pooling to occur more naturally than in tasks that require some selective attention. For example, Ingvalson and Wenger (2005) used a change detection task, much like the composite face task, to demonstrate supercapacity with faces. With the caveat that analysis of capacity requires an assumption of context invariance, which may not hold experimentally, we feel that this analysis has much explanatory power for the present experiments. Importantly, it suggests a way forward for reconciling the composite face task with tasks aimed at uncovering the underlying mental architecture of face processing. Although our task can clearly uncover coactivity when it exists (see also, Little et al., 2013), our tasks may have been susceptible to subject specific selective attention strategies.

What underlies differences in performance with upright and inverted, aligned and misaligned faces in the composite face task? The absence of strong evidence for coactivity in our task implicates some other mechanism for explaining the interaction between upright faces and inverted or misaligned faces which arises in the composite face task. Our best guess for this mechanism is a failure of selective attention (Chua et al., 2014, 2015; Richler et al., 2011; Richler et al., 2011). A complete explanation of our results and the composite face task results is that processing in the composite face task may proceed in parallel or in serial (in most cases), but is coupled with an inability to completely ignore one face half when faces are presented in an upright and aligned fashion. By contrast, inversion and misalignment aid selective attention to the relevant face half.

One difference between our task and the composite face paradigm is the analysis of individuals versus group averaged data, respectively. There are a number of arguments for why analysing group data can be misleading. For instance, the average data may not accurately represent any single individual (e.g., Estes & Maddox, 2005; Liew, Howe, & Little, 2016; Navarro, Griffiths, Steyvers, & Lee, 2006). Hence, averaging in the composite face paradigm may result in masking individual differences in expression of the interaction effect (between congruency and alignment in the composite design). While there is some evidence for a correlation between different holistic processing tasks (DeGutis, Wilmer, Mercado, & Cohan, 2013; but see Wang, Li, Fang, Tian, & Liu, 2012), there is debate about the nature of these relationships (Richler, Floyd, & Gauthier, 2015; Sunday, Richler, & Gauthier, in press). One path to reconciling the differing results between tasks would be to conduct individual level analyses on a double factorial recognition task in which the processing requirements are much closer to the composite face task but which has the ability to draw inferences as we do in our present experiments. We leave this as a target for future research.

References

Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of response times: Identification of cognitive mechanisms through mathematical modeling. *The Oxford Handbook of Computational and Mathematical Psychology*, 63-98.

Algom, D., & Fitousi, D. (2016). Half a century of research on Garner interference and the Separability-integrality distinction. *Psychological Bulletin, 142*(12), 1352-1383.

Algom, D., Fitousi, D., & Eidels, A. (2017). Bridge-building: SFT interrogation of major cognitive phenomena. In D. R. Little, N. Altieri, M. Fific & C-T. Yang (Eds.). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. London: Elsevier.

Altieri, N., Fific, M., Little, D. R. & Yang, C-T. (2017). Historical foundations and a tutorial introduction to Systems Factorial Technology. In D. R. Little, N. Altieri, M. Fific & C-T. Yang (Eds.). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. London: Elsevier.

Amishav, R., & Kimchi, R. (2010). Perceptual integrality of componential and configural information in faces. *Psychonomic Bulletin & Review*, *17*(5), 743-748.

Ashby, F. G., Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 33-53.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*(2), 154.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*, 281–316.

Blaha, L. (2017). An examination of the task demands on the elicited processing capacity. In D. R. Little, N. Altieri, M. Fific & C-T. Yang (Eds.). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. London: Elsevier.

Blunden, A. G., Wang, T., Griffiths, D. W., Little, D. R. (2015). Logical-rules and the classification of integral dimensions: Individual differences in the processing of arbitrary dimensions. *Frontiers in Psychology, 5*(Article 1531), 1-24.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153-178.

Chance, J. E., & Goldstein, A. G. (1981). Depth of processing in response to own- and other-race faces. *Personality and Social Psychology Bulletin, 7,* 475-480.

Cheng, X. J., Moneer, S., Christie, N., & Little, D. R. (2017). Categorization, Capacity, and Resilience. In D. R. Little, N. Altieri, M. Fifić & C-T. Yang (Eds.). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. London: Elsevier.

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1327-1336.

Chua, K. W., Richler, J. J., & Gauthier, I. (2014). Becoming a Lunari or Taiyo expert: Learned attention to parts drives holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1174.

Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *Journal of Experimental Psychology: General, 144,* 723.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126,* 87–100

Donnelly, N., Cornes, K., & Menneer, T. (2012). An examination of the processing capacity of features in the Thatcher illusion. *Attention, Perception, & Psychophysics, 74*(7), 1475-1487.

Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology, 55*(2), 176-190.

Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychonomic Bulletin & Review, 12,* 403–408.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628-634.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482-498.

Fifić, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review, 117(2),* 309–348.

Fifić, M., Nosofsky, R. M., & Townsend, J. T. (2008). Information-processing architectures in multidimensional classification: A validation test of the systems factorial technology. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 356–375.

Fifić, M., & Townsend, J. T. (2010). Information-processing alternatives to holistic perception: Identifying the mechanisms of secondary-level holism within a categorization paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36*(5), 1290–1313.

Fitousi, D. (2015). Composite faces are not processed holistically: Evidence from the Garner and redundant target paradigms. *Attention, Perception, & Psychophysics, 77*(6), 2037-2060.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103,* 322-330.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, *24*(6), 997-1016.

Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Sciences*, *23,* 427–434.

Goldstone, R. L., & Steyvers, M. (2001). The sensitisation and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General, 130*(1), 116-139.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Griffiths, D. W., Blunden, A. G., & Little, D. R. (2017). Logical-rule based models of categorization: Using Systems Factorial Technology to understand feature and dimensional processing. In D. R. Little, N. Altieri, M. Fifić & C-T. Yang (Eds.). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. London: Elsevier.

Harding, B., Goulet, M-A., Jolin, S., Tremblay, C., Villeneuve, S-P., & Guillaume, D. (2016). Systems factorial technology explained to humans. *The Quantitative Methods for Psychology, 12,* 39-57.

Hayward, W. G., Crookes, K., Chu, M. H., Favelle, S. K., & Rhodes, G. (2016). Holistic processing of face configurations and components. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 1482-1489.

Heathcote, A., Brown, S. D., & Wagenmakers, E-J. (2015). An introduction to good practices in cognitive modelling. In B. U. Forstmann & E-J. Wagenmakers. (Eds.). *An Introduction to Model-based Cognitive Neuroscience.* New York: Springer.

Holmes, W. R. (2015). A practical guide to the probability density approximation (PDA) with improved implementation and error characterization. *Journal of Mathematical Psychology, 68-69,* 13-24.

Houpt, J. W., & Townsend, J. T. (2011). An extension of SIC predictions to the wiener coactive model. *Journal of Mathematical Psychology, 55*(3), 267-270.

Ingvalson, E. M., & Wenger, M. J. (2005). A strong test of the dual-mode hypothesis. *Perception & Psychophysics, 67*(1), 14-35.

Kantowitz, B. H. (1974). *Human information processing: Tutorials in performance and cognition.* Hillsdale, NJ: Erlbaum.

Kayser, A. (1997). *Heads*. New York, NY: Abbeville Publishing Group.

Li, X., Liang, Z., Kleiner, M., & Lu, Z. L. (2010). RTbox: A device for highly accurate response time measurements. *Behavioural Research Methods, 42*, 212-225.

Liew, S. X., Howe, P. D., & Little, D. R. (2016). The appropriacy of averaging in the study of context effects. *Psychonomic Bulletin & Review*, *23*, 1639-1646.

Little, D. R., Altieri, N., Fifić, M. & Yang, C-T. (2017). *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms.* London: Elsevier.

Little, D. R., Eidels, A., Fifić, M., & Wang, T. (2015). Understanding the influence of distractors on workload capacity. *Journal of Mathematical Psychology, 68,* 25-36.

Little, D. R., Nosofsky, R. M., & Denton, S. E. (2011). Response-time tests of logical-rule models of categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37*(1), 1–27.

Little, D. R., Nosofsky, R. M., Donkin, C., & Denton, S. E. (2013). Logical rules and the classification of integral-dimension stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*(3), 801–820.

Little, D. R., Wang, T., & Nosofsky, R. M. (2016). Sequence-sensitive exemplar and decision-bound accounts of speeded-classification performance in a modified Garner-tasks paradigm. *Cognitive Psychology, 89*, 1-38.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press.

Maurer, D., LeGrand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

Melara, R. D., Marks, L. E., & Lesko, K. E. (1992). Optional processes in similarity judgments. *Attention, Perception, & Psychophysics, 51*(2), 123-133.

Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*(5), 627-645.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247-279.

Moneer, S., Wang, T., & Little, D. R. (2016). The Processing Architectures of Whole-Object Features: A Logical-Rules Approach. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 1443-1465.

Navarro, D. J., Griffiths, T. L., Steyvers, M., & Lee, M. D. (2006). Modeling individual differences using Dirichlet processes. *Journal of Mathematical Psychology, 50,* 101–122.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology, 43,* 25-53.

Perry, L., Blaha, L., & Townsend, J. (2008). Reassessing the architecture of same-different face judgments. *Journal of Vision*, *8*(6), 889-889.

Pomerantz, J.R., & Pristach, E.A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635–649.

Raab, D. H. (1962). Division of Psychology: Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24*(5 Series II), 574-590.

Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review, 85*(2), 59-108.

Richler, J.J., Floyd, R.J. & Gauthier, I. (2015). About-face on face recognition ability and holistic processing, *Journal of Vision*, 15(9):15

Richler, J. J., & Gauthier, I. (2013). When intuition fails to align with data: A reply to Rossion (2013). *Visual Cognition, 21*(2), 254-276.

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research, 51*(3), 333-342.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, Mechanisms, and Measures of Holistic Processing. *Frontiers in Psychology, 3*(Article 553), 1-6.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2015). Holistic processing does not require configural variability. *Psychonomic Bulletin & Review, 22*(4), 974-979.

Richler, J. J., Tanaka, J. W., Brown, D. D., & Gauthier, I. (2008). Why does selective attention to parts fail in face processing*? Journal of Experimental Psychology: Learning, Memory and Cognition, 34*(6), 1356–1368.

Richler, J.J., Gauthier, I., Wenger, M., & Palmeri, T.J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 328-342.

Richler, J. J., Wong, Y. K. & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Sciences, 20*, 129–134.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274-289.

Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision, 8(4),* 1–13.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21(6),* 329-330.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, Search, and Attention. *Psychological Review, 84,* 1-65.

Schwaninger, A., & Mast, F. W. (2005). The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism. *Japanese Psychological Research, 47*(3), 216-222.

Schweickert, R. (1992). Information, time, and the structure of mental events: A twenty-five year review. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance: Vol. 14. Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience—A silver jubilee* (pp. 535–566). Cambridge, MA: MIT Press.

Shen, J., & Palmeri, T. J. (2015). The perception of a face can be greater than the sum of its parts. *Psychonomic Bulletin & Review, 22*, 710-716.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 4,* 421–457.

Steyvers, M. (1999). Morphing techniques for generating and manipulating face images. *Behaviour Research Methods, Instruments, & Computers, 31*, 359-369.

Sunday, M.A., Richler, J.J. & Gauthier, I. (in press). Limited evidence of individual differences in holistic processing in different versions of the part-whole paradigm. *Attention, Perception & Psychophysics*.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 46*(2), 225–245.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9,* 483-484.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28,* 363–400.

Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin, 108*(3), 551.

Townsend, J.T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review, 18*(4), 659–681.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39,* 321-359.

Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology, 56*, 476-494.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review, 111*(4), 1003.

Turner, B. M., & Sederberg, P. B. (2014). A generalized, likelihood-free method for posterior estimation. *Psychonomic Bulletin & Review*, *21*(2), 227-250.

Turner, B. M., Sederberg, P. B., Brown, S. D., & Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods, 18*(3), 368.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*, 169–177.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research, 11*, 3571–3594.

Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 32*(3), 755.

Wong, Y. K. & Gauthier, I. (2010). Holistic processing of musical notation: Dissociating failures of attention in experts and novices. *Cognitive and Affective Behavioral Neuroscience, 10*, 541–551.

Yang, C-T., Fifić, M., Chang, T-Y., & Little, D. R. (2017). Systems factorial technology provides new insights on the other-race effect. *Psychonomic Bulletin & Review.* Advance online publication. doi:10.3758/s13423-017-1305-9

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Young, A. M., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception, 10,* 747-759.

Author Note

This work was supported by ARC Discovery Project Grant DP160102360 and a Melbourne Research Grant Support Scheme Grant to Daniel R. Little, and NSF Grant SBE-1257098 and Vanderbilt International Office Grant to Thomas J. Palmeri. Thanks to Adam Osth for helpful discussions about DE-MCMC modelling. We thank Daniele Martinie for assistance with data collection. Portions of this work were completed as part of the fulfilment of honours degree projects by Xue Jun Cheng and Callum McCarthy.

Correspondence concerning this article should be addressed to Xue Jun Cheng, School of Psychological Sciences, University of Melbourne, Melbourne, Victoria 3010, Australia. E-mail: xjcheng@student.unimelb.edu.au or to Daniel R. Little School of Psychological Sciences, University of Melbourne, Melbourne, Victoria 3010, Australia. E-mail: daniel.little@unimelb.edu.au

Appendix A

Multidimensional Scaling

MDS solutions were obtained from participants through a Human Intelligence Task (HIT) on the Amazon Mechanical Turk platform. Participants completed the task for either the upright aligned (N = 31), upright misaligned (N = 32), inverted aligned (N = 29), or inverted misaligned (N = 31) conditions. For each condition, there were 36 unique pairings of the 9 stimuli. For half of the participants, each pair was presented six times for each subject; for the remaining half, each pair was presented twice. The order of presentation was completely randomised, as was the left-right presentation of each face. The experiment was self-paced.

We computed the averaged similarity rating for each pair of stimuli and found the two-dimensional scaling solutions for each condition by fitting these averaged ratings using a model which assumed a negative linear relationship between the distance between the estimated coordinates and the predicted similarity ratings. To find the best fitting coordinates, we minimized the sum-of-squared deviations between the predicted and observer ratings from 100 starting points chosen to span the coordinate space. There were 20 parameters in total (the 18 coordinate values and the slope and the intercept of the negative linear distance-to-similarity function) used to fit the 36 similarity ratings. We fitted a constrained version of the model in which the nine items were fixed to fall on a 3×3 grid but variation was allowed between each X and Y value.

The full model provides excellent fits of the data; accounting for 98-99% of the variance. The constrained model also provides similarly good fits by accounting for 95-98% of the variance. Consequently, the constrained model is the preferred solution since it provides an equally good fit of the data as the full model but using fewer parameters.

The MDS solutions for the upright conditions align closely to our morph settings (see Figure A1). As can be seen in Figure A2, the MDS solutions for the inverted faces do not match our physical manipulations of the stimuli (i.e., equal space between each of the 9 stimuli on both dimensions as they are 0%, 50%, 100% morphed combinations). While the three levels on the Bottom Half dimension are relatively equally-spaced, two levels in the Top Half dimension are harder to discriminate from the third level. This caused concerns as to whether it would be appropriate to use these solutions in model-fitting. As such, we decided to obtain individual MDS solutions for each of our eight participants. Though not presented, all eight participants in this study revealed very similar MDS spaces to the ones in Figure A2; hence, we continued model-fitting with these MDS solutions.

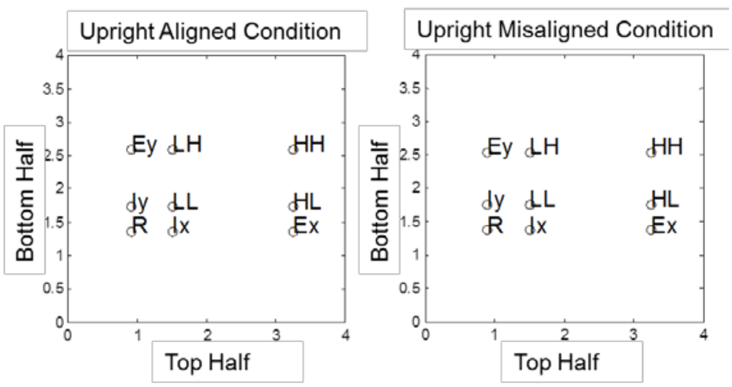


Figure A1. MDS solutions for the upright aligned (left) and misaligned (right) conditions.

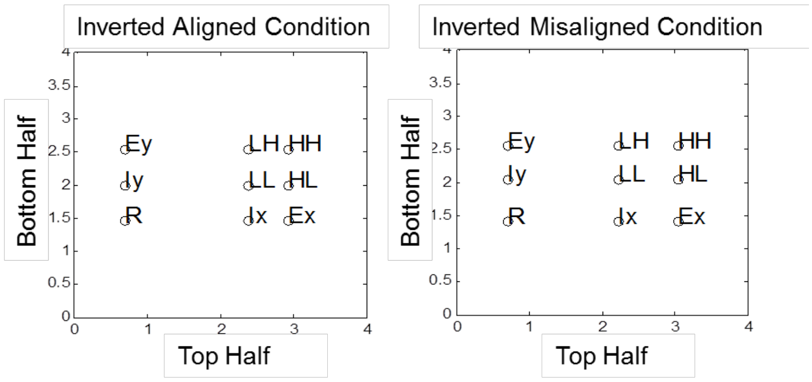


Figure A2. MDS solutions for the inverted aligned (left) and misaligned (right) conditions.

In the above analyses, we assumed that the composite faces are separable stimuli and used city block distance in computing the distance between each stimulus. However, it is plausible that these faces are processed holistically and thus the scaling solution with a Euclidean distance may provide a better fit. To examine this, we allowed the distance metric, *r*, to vary freely while estimating the best-fitting coordinates. A summary of our model fits is presented in Table A1. Unsurprisingly, with the additional parameter, this model provides the best fit of the scaling solution with the full model accounting for 99% of the variance and the constrained model accounting for 95-99% of the variance. More importantly, *r* for each of the face sets (i.e., for each condition) is close to one, confirming the assumption that these faces are separable stimuli.

Table A1

Summary of the Minkowski Model Fits

	Full Model				Constrained Model			
Condition	r	SSD	R^2	BIC	r	SSD	R^2	BIC
UA	1.33	0.76	.99	-84.86	1.30	1.90	.97	-102.01
UM	1.16	0.42	.99	-107.75	1.20	0.81	.99	-134.46
IA	1.10	0.80	.99	-78.08	1.10	2.67	.95	-84.83
IM	0.93	0.47	.99	-99.85	1.07	0.90	.99	-126.84

Note. UA = Upright Aligned, UM = Upright Misaligned, IA = Inverted Aligned, IM = Inverted Misaligned

Appendix B

Table B1

Prior parameter values and transformations used in DE-MCMC sampling.

Parameter	Transformation	Distribution	Prior Parameter Values	
D_{top}	$\widehat{D}_{Top} = logit\{(D_{Top} - x_0)/(x_1 - x_0)\}$	Normal	μ= 0	σ= .5
D_{bot}	$\widehat{D}_{Bot} = logit\{(D_{Bot} - y_0)/(y_1 - y_0)\}$	Normal	μ= 0	σ= .5
σ_{Top}	$\hat{\sigma}_{Top} = log(\sigma_{Top})$	Normal	μ= -1.5	σ= .2
σ_{Bot}	$\hat{\sigma}_{Bot} = log(\sigma_{Bot})$	Normal	μ= -1.5	σ= .2
A	$\hat{A} = log(A)$	Normal	μ= -1.05	σ= .2
b_A	$\hat{b}_A = log(b_A - A)$	Normal	μ= -1.05	σ= 1
b_B	$\hat{b}_B = log(b_B - A)$	Normal	μ= -1.05	σ= 1
s	$\hat{s} = log(s)$	Normal	μ= -1.39	σ= .5
t_0	$\hat{t}_0 = log(t_0)$	Normal	μ= -1.51	σ= .2
p_X	$\hat{p}_x = logit(p_x)$	Normal	μ= 0	σ= 2
p_{Serial}	$\hat{p}_{Serial} = logit(p_{Serial})$	Normal	μ= 0	σ= 2
m	$\widehat{m} = log(m)$	Normal	μ= 0	σ= 2
v	$\hat{v}_x = logit(v_x)$	Normal	μ= 0	σ= 1

Note: Each parameter was transformed to the range -∞ to +∞. All parameters were then assigned Normal distribution priors.

Appendix C

Composite Face Task Results

We conducted a composite face task experiment using the complete experimental design described by Richler and Gauthier (2013). In this task, a composite face is presented on each trial with one face half (the top or the bottom) cued. After a brief interval, a test face is presented and subjects must respond as to whether the cued face part is the same or different. On *congruent* trials, if the cued face half is the same (or different), the other face half is also the same (or different). On *incongruent* trials, when the cued face half is same (or different), the other face half is different (or the same). The key analysis concerns comparison between when the study face is aligned or misaligned. If processing is holistic (if the subject is unable to ignore the uncued face half), then we should find that the d-prime value computed from the hit and false alarm rates (Green & Swets, 1966) is higher for congruent trials than for incongruent trials *and* that this difference is reduced when the faces are misaligned. Holistic processing is therefore inferred from an interaction between congruency and alignment.

In the present task, we tested all four conditions from our categorization experiment: Upright Aligned, Upright Inverted, Inverted Aligned, and Inverted Misaligned. Forty participants were recruited from the Melbourne School of Psychological Sciences' Research Experience Program. There were 16 blocks of 48 trials. Only the four corner faces from Figure 1 (middle panel) were presented. At the beginning of each block, participants were cued to attend to either the top or bottom face half. The study face was always upright and aligned, the test face could be either upright or inverted or aligned or misaligned. The test face direction and alignment was randomized on each trial such that all combinations of study and test faces were presented across the course of the experiment. The 256 (2 Face Set - upright vs inverted x 2 Cued Half x 2 Same vs Diff x 2 Congruent vs Incongruent x 2 Upright vs Inverted x 2 Aligned vs Split x 4 Corner Face) were repeated three times each for a total of 768 trials.

The main results are shown in Figure C1. We analysed these results using a 2 Direction (Upright vs Inverted) x 2 Alignment (Aligned vs Split) x 2 Congruency Repeated Measures ANOVA. These results are shown in Table C1. The most interesting results are shown most clearly in the planned comparison repeated measures ANOVAs on the d-prime measure (but also reflected in the three-way ANOVA). These results show that there is a congruency x alignment interaction for the upright faces, $F(1, 39) = 6.53$, $MSE = 0.14$, $p = .01$, *and* for the inverted faces, $F(1, 39) = 4.63$, $MSE = 0.12$, $p = .04$ (similar results were found by Richler et al. 2011). The congruency x direction interaction was not significant for the aligned faces, $F(1, 39) = 3.41$, $MSE = 0.18$, $p > .07$, or for the misaligned faces $F(1, 39) = 1.83$, $MSE = 0.17$, $p = .18$. A good summary would be that, for d-prime, there is a main effect of direction (upright vs inverted) and that there is a congruency x alignment interaction such that overall upright faces are easier to recognize than inverted faces, but the hallmark of holistic processing appears for both upright and inverted faces. That is, misaligning the face disrupts holistic processing, but inverting the face just dampens performance altogether (as observed by Richler et al., 2011).

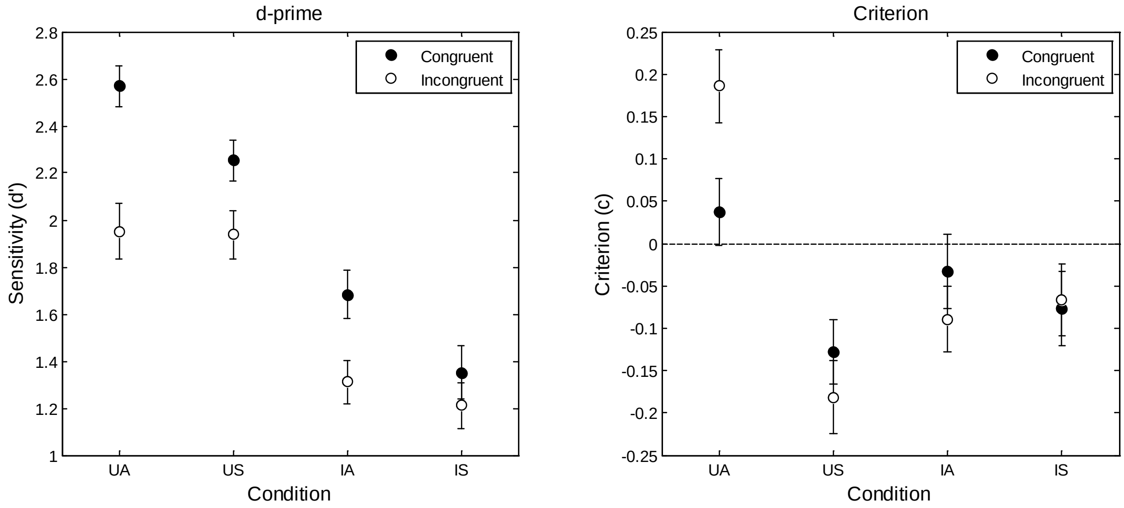


Figure C1. Sensitivity and criterion results for Upright Aligned, Upright Misaligned, Inverted Aligned, and Inverted Misaligned Conditions. (Standard errors are shown).

Table C1

Repeated Measures Direction x Alignment x Congruency ANOVA

Effect	df Between	df Within	MSE	F	p
Direction (Upright vs Inverted)	1	39	0.20	243.85	< .001
Alignment (Aligned vs Split)	1	39	0.17	16.78	< .001
Congruency	1	39	0.20	51.93	< .001
Direction x Alignment	1	39	0.17	0.32	> .05
Direction x Congruency	1	39	0.19	4.73	< .05
Alignment x Congruency	1	39	0.10	14.45	< .001
Dir. x Align. x Cong.	1	39	0.16	0.16	> .05

The criterion estimates, though of less interest in the current analysis, seem to be in the same direction for all conditions except for the upright aligned condition and are reported for completeness (see Table C2).

Table C2

Repeated Measures Direction x Alignment x Congruency ANOVA

Effect	df Between	df Within	MSE	F	p
Direction (Upright vs Inverted)	1	39	0.11	1.54	> .05
Alignment (Aligned vs Split)	1	39	0.06	24.36	< .001
Congruency	1	39	0.03	0.36	> .05
Direction x Alignment	1	39	0.06	21.97	< .001
Direction x Congruency	1	39	0.04	2.4	> .05
Alignment x Congruency	1	39	0.04	2.38	> .05
Dir. x Align. x Cong.	1	39	0.04	9.09	< .01

Table 1

Summary of Predictions for Theories of Composite Face Processing

		Target Category Predictions			
Theory	Condition	MIC	SIC	Contrast Category	Inference
Holistic processing of upright faces only	Upright Aligned	Overadditivity (MIC > 0)	Mostly positive	Interior < Exterior	Coactive
	Other Conditions	Underadditive or additive (MIC ≤ 0)	Negative or S-shaped	Interior ≥ Exterior	Serial or Parallel
Composite faces processed as separable objects	All Conditions	Underadditive or additive (MIC ≤ 0)	Negative or S-shaped	Interior ≥ Exterior	Serial or Parallel
Holistic processing of aligned faces	Aligned	Overadditivity (MIC > 0)	Mostly positive	Interior < Exterior	Coactive
	Misaligned	Underadditive or additive (MIC ≤ 0)	Negative or S-shaped	Interior ≥ Exterior	Serial or Parallel

Table 2

Observed Mean Correct and Error RTs (ms) and Error Rates for Individual Stimuli for Each Observer in the Upright Face Conditions

		Item								
Observer	Variable	HH	HL	LH	LL	Ex	Ix	Ey	Iy	R
UA1	RT correct	950	1292	1201	1473	1319	1281	1087	1110	1030
	RT error	-	1916	1424	1648	1446	1594	1282	1305	1110
	p(error)	0.00	0.10	0.14	0.27	0.11	0.07	0.05	0.03	0.01
UA2	RT correct	796	1041	1069	1264	989	1034	918	1009	842
	RT error	1026	1152	1264	1341	1267	1447	940	1772	1362
	p(error)	0.01	0.08	0.07	0.14	0.15	0.07	0.04	0.06	0.01
UA3	RT correct	683	1020	887	1154	1106	1167	1014	907	745
	RT error	-	1503	1903	1544	1626	1825	1163	1206	1537
	p(error)	0.00	0.04	0.01	0.07	0.09	0.09	0.06	0.02	0.01
UA4	RT correct	888	1256	1339	1656	1104	1260	1210	1079	1007
	RT error	-	1595	1730	1623	1763	1741	1838	2287	1200
	p(error)	0.00	0.01	0.04	0.17	0.08	0.05	0.05	0.01	0.00
UM1	RT correct	914	1196	1290	1492	1166	1172	1064	1068	995
	RT error	-	1287	1253	1584	1527	1674	1440	1653	2156
	p(error)	0.00	0.02	0.06	0.10	0.13	0.05	0.03	0.01	0.00
UM2	RT correct	822	970	1112	1224	925	982	976	974	884
	RT error	1237	2068	1385	1435	1265	1299	1616	1803	1518
	p(error)	0.00	0.02	0.06	0.17	0.10	0.04	0.13	0.09	0.00
UM3	RT correct	783	989	1219	1353	1061	1103	987	938	867
	RT error	824	1297	1219	1138	1376	1570	1538	1669	1432
	p(error)	0.01	0.03	0.04	0.13	0.14	0.11	0.05	0.05	0.01
UM4	RT correct	684	857	829	934	955	880	724	735	686
	RT error	718	1202	1010	954	1054	1200	945	930	987
	p(error)	0.01	0.09	0.11	0.22	0.23	0.21	0.04	0.03	0.01

Note: UA - Upright Aligned, UM - Upright Misaligned; HH = high-high stimulus x_2y_2; HL = high-low stimulus x_2y_1; LH = low-high stimulus x_1y_2; LL = low-low stimulus x_1y_1; R = redundant stimulus x_0y_0; Ix and Iy denote the interior stimuli on the Top Half x_1y_0 and Bottom Half x_0y_1 dimensions respectively; Ex and Ey denote the exterior stimuli on the Top Half x_2y_0 and Bottom Half x_0y_2 dimensions respectively.

Table 3

Observed Mean Correct and Error RTs (ms) and Error Rates for Individual Stimuli for Each Observer in the Inverted Face Conditions

		Item								
Observer	Variable	HH	HL	LH	LL	Ex	Ix	Ey	Iy	R
IA1	RT correct	668	821	704	885	892	846	711	684	645
	RT error	-	963	1110	1036	1030	1230	603	1317	-
	p(error)	0.00	0.02	0.03	0.10	0.11	0.02	0.03	0.01	0.00
IA2	RT correct	806	930	896	993	889	912	769	745	698
	RT error	1062	1009	1061	1036	835	1105	736	917	757
	p(error)	0.03	0.14	0.06	0.17	0.08	0.04	0.05	0.01	0.01
IA3	RT correct	1257	1409	1448	1591	1514	1469	1101	1095	1022
	RT error	2059	1902	1827	2013	1812	1847	1316	1615	931
	p(error)	0.01	0.05	0.02	0.13	0.17	0.13	0.02	0.03	0.01
IA4	RT correct	713	753	753	829	840	820	600	624	624
	RT error	923	997	896	881	868	1049	597	790	-
	p(error)	0.00	0.02	0.01	0.03	0.07	0.05	0.01	0.01	0.00
IA5	RT correct	691	835	825	1057	937	888	761	732	641
	RT error	899	1377	1073	1216	1604	1517	874	0	1697
	p(error)	0.00	0.05	0.01	0.12	0.06	0.03	0.02	0.00	0.00
IM1	RT correct	704	773	730	817	680	673	652	661	621
	RT error	462	900	695	916	861	721	1274	1131	809
	p(error)	0.00	0.01	0.02	0.03	0.04	0.03	0.01	0.01	0.00
IM2	RT correct	745	819	817	910	953	903	576	584	589
	RT error	850	549	531	889	1033	1075	-	-	1319
	p(error)	0.01	0.00	0.02	0.04	0.04	0.02	0.00	0.00	0.00
IM3	RT correct	899	1035	1039	1180	1187	1310	834	813	838
	RT error	-	2364	1557	2179	1465	2053	1342	2736	-
	p(error)	0.00	0.01	0.03	0.06	0.05	0.04	0.01	0.01	0.00
IM4	RT correct	751	841	822	932	811	821	740	740	682
	RT error	602	822	1084	1340	1277	1835	1356	1242	157
	p(error)	0.00	0.03	0.02	0.02	0.04	0.03	0.01	0.01	0.00

Note: IA - Inverted Aligned, IM - Inverted Misaligned; HH = high-high stimulus x_2y_2; HL = high-low stimulus x_2y_1; LH = low-high stimulus x_1y_2; LL = low-low stimulus x_1y_1; R = redundant stimulus x_0y_0; Ix and Iy denote the interior stimuli on the Top Half x_1y_0 and Bottom Half x_0y_1 dimensions respectively; Ex and Ey denote the exterior stimuli on the Top Half x_2y_0 and Bottom Half x_0y_2 dimensions respectively.

Table 4

Target Category Statistical Results for the Individual Participants in the Upright Conditions

Variable	*df*	*F*	*p*	*df*	*F*	*p*
	Participant UA1			Participant UM1		
Session	6	19.17	.00	6	26.55	.00
Top	1	136.67	.00	1	156.13	.00
Bottom	1	66.97	.00	1	310.55	.00
Session x T	6	2.04	.06	6	1.16	.33
Session x B	6	1.29	.26	6	2.70	.01
Top x Bottom	1	2.22	.14	1	4.98	.03
Session x T x B	6	0.68	.66	6	1.18	.31
Error	1192			1336		
	Participant UA2			Participant UM2		
Session	6	17.04	.00	6	182.32	.00
Top	1	95.44	.00	1	104.45	.00
Bottom	1	109.91	.00	1	393.25	.00
Session x T	6	1.19	.31	6	3.10	.01
Session x B	6	2.19	.04	6	16.33	.00
Top x Bottom	1	1.43	.23	1	0.47	.49
Session x T x B	6	0.98	.44	6	1.30	.26
Error	1264			1307		
	Participant UA3			Participant UM3		
Session	6	7.88	.00	6	59.80	.00
Top	1	365.38	.00	1	117.87	.00
Bottom	1	114.15	.00	1	619.07	.00
Session x T	6	3.02	.01	6	1.07	.38
Session x B	6	1.08	.37	6	4.38	.00
Top x Bottom	1	4.83	.03	1	3.67	.06
Session x T x B	6	1.52	.17	6	0.65	.69
Error	1355			1327		
	Participant UA4			Participant UM4		
Session	6	20.56	.00	5	16.37	.00
Top	1	206.21	.00	1	174.29	.00
Bottom	1	322.35	.00	1	95.19	.00
Session x T	6	1.69	.12	5	0.52	.76
Session x B	6	2.94	.01	5	1.45	.20
Top x Bottom	1	1.33	.25	1	5.31	.02
Session x T x B	6	2.04	.06	5	2.04	.07
Error	1315			1061		

Note: UA = Upright Aligned, UM = Upright Misaligned, T = Top, B = Bottom.

Table 5

Contrast Category Statistics for the Upright Conditions

Stimulus pair	M_{diff}	*t*	*df*	*p*	M_{diff}	*t*	*df*	*p*
	UA1				UM1			
E_{Top} - I_{Top}	-23	-0.66	665	.51	-4	-0.14	681	.89
E_{Bottom} - I_{Bottom}	40	1.04	633	.30	-6	-0.17	636	.87
E_{Top} - R	57	1.81	675	.07	69	2.19	685	.03
I_{Top} - R	80	2.39	682	.02	73	2.30	692	.02
E_{Bottom} - R	292	8.39	657	< .001	172	5.20	651	< .001
I_{Bottom} - R	252	7.27	668	< .001	177	5.30	681	< .001
	UA2				UM2			
E_{Top} - I_{Top}	-91	-2.78	661	.01	2	0.05	617	.96
E_{Bottom} - I_{Bottom}	-45	-1.45	620	.15	-57	-2.01	646	.04
E_{Top} - R	75	2.99	678	< .001	92	2.81	649	.01
I_{Top} - R	166	5.50	671	< .001	90	2.75	664	.01
E_{Bottom} - R	147	5.66	641	< .001	41	1.45	660	.15
I_{Bottom} - R	192	7.02	667	< .001	98	3.22	682	< .001
	UA3				UM3			
E_{Top} - I_{Top}	107	4.27	672	< .001	50	1.54	664	.12
E_{Bottom} - I_{Bottom}	-60	-1.58	632	.11	-42	-1.32	608	.19
E_{Top} - R	269	12.53	676	< .001	120	4.08	679	< .001
I_{Top} - R	162	8.11	690	< .001	71	2.47	679	.01
E_{Bottom} - R	361	13.69	662	< .001	194	6.96	644	< .001
I_{Bottom} - R	422	14.11	664	< .001	236	8.03	658	< .001
	UA4				UM4			
E_{Top} - I_{Top}	131	3.66	675	< .001	-31	-1. 81	573	.07
E_{Bottom} - I_{Bottom}	-156	-4.16	652	< .001	59	2.95	492	.003
E_{Top} - R	202	5.72	679	< .001	41	2.99	581	.003
I_{Top} - R	72	2.04	692	.04	72	4.58	586	<.001
E_{Bottom} - R	96	2.74	668	.01	268	16.65	531	< .001
I_{Bottom} - R	253	6.79	680	< .001	209	13.08	555	< .001

Note: UA = Upright Aligned, UM = Upright Misaligned. E = Exterior, I = Interior, R = Redundant.

Table 6

Target Category Statistical Results for the Inverted Conditions

Variable	*df*	*F*	*p*	*df*	*F*	*p*
	Participant IA1			Participant IM1		
Session	6	22.18	.00	6	82.98	.00
Top	1	217.46	.00	1	56.41	.00
Bottom	1	23.52	.00	1	10.98	.00
Session x T	6	1.43	.20	6	2.53	.02
Session x B	6	2.26	.04	6	0.24	.96
Top x Bottom	1	2.42	.12	1	0.84	.36
Session x T x B	6	0.46	.83	6	0.49	.81
Error	1322			1348		
	Participant IA2			Participant IM2		
Session	6	49.69	.00	6	25.46	.00
Top	1	52.37	.00	1	49.57	.00
Bottom	1	26.01	.00	1	45.38	.00
Session x T	6	1.44	.20	6	1.61	.14
Session x B	6	2.10	.05	6	1.89	.08
Top x Bottom	1	0.69	.41	1	0.72	.40
Session x T x B	6	1.30	.25	6	0.83	.54
Error	1233			1346		
	Participant IA3			Participant IM3		
Session	6	17.35	.00	6	78.91	.00
Top	1	18.05	.00	1	58.02	.00
Bottom	1	26.81	.00	1	69.38	.00
Session x T	6	1.15	.33	6	2.75	.01
Session x B	6	4.63	.00	6	10.50	.00
Top x Bottom	1	0.00	1.00	1	0.68	.41
Session x T x B	6	2.06	.06	6	0.84	.54
Error	1292			1337		
	Participant IA4			Participant IM4		
Session	6	7.71	.00	6	14.71	.00
Top	1	29.41	.00	1	26.36	.00
Bottom	1	29.60	.00	1	17.43	.00
Session x T	6	0.83	.54	6	1.21	.30
Session x B	6	2.18	.04	6	0.92	.48
Top x Bottom	1	2.82	.09	1	0.29	.59
Session x T x B	6	1.03	.40	6	1.00	.42
Error	1348			1347		
	Participant IA5					
Session	6	27.76	.00			
Top	1	152.13	.00			
Bottom	1	136.21	.00			
Session x T	6	0.95	.46			
Session x B	6	3.78	.00			
Top x Bottom	1	9.56	.00			
Session x T x B	6	0.89	.50			
Error	1306					

Note: IA = Inverted Aligned, IM = Inverted Misaligned, T = Top, B = Bottom.

Table 7

Contrast Category Statistics for the Inverted Conditions

Stimulus pair	*M*	*t*	*df*	*p*	*M*	*t*	*df*	*p*
	IA1				IM1			
E_{Top} - I_{Top}	27.18	2.04	685	0.04	-9.15	-0.54	692	0.59
E_{Bottom} - I_{Bottom}	46.23	2.39	648	0.02	6.86	0.43	675	0.67
E_{Top} - R	65.68	5.48	687	0.00	30.58	2.06	694	0.04
I_{Top} - R	38.50	3.03	696	0.00	39.73	2.49	694	0.01
E_{Bottom} - R	247.34	15.01	657	0.00	58.10	4.08	683	0.00
I_{Bottom} - R	201.11	13.70	689	0.00	51.24	3.33	688	0.00
	IA2				IM2			
E_{Top} - I_{Top}	23.60	1.31	678	0.19	-8.75	-0.61	698	0.54
E_{Bottom} - I_{Bottom}	-22.31	-1.07	655	0.28	50.38	2.84	675	0.00
E_{Top} - R	70.34	4.05	677	0.00	-13.42	-1.04	697	0.30
I_{Top} - R	46.75	2.72	691	0.01	-4.67	-0.32	697	0.75
E_{Bottom} - R	190.76	10.28	667	0.00	363.97	23.07	681	0.00
I_{Bottom} - R	213.06	11.37	678	0.00	313.59	20.14	690	0.00
	IA3				IM3			
E_{Top} - I_{Top}	5.3425	0.13	680	0.90	21.09	0.83	692	0.41
E_{Bottom} - I_{Bottom}	45.197	0.85	586	0.40	-123.60	-3.00	660	0.00
E_{Top} - R	79.186	2.03	687	0.04	-3.77	-0.13	694	0.89
I_{Top} - R	73.843	1.93	687	0.05	-24.85	-0.93	696	0.35
E_{Bottom} - R	492.54	11.34	635	0.00	348.94	10.51	678	0.00
I_{Bottom} - R	447.34	10.09	645	0.00	472.54	12.56	680	0.00
	IA4				IM4			
E_{Top} - I_{Top}	-24.48	-2.31	690	0.02	-0.20	-0.01	691	0.99
E_{Bottom} - I_{Bottom}	19.84	1.28	656	0.20	-10.59	-0.43	672	0.67
E_{Top} - R	-24.38	-2.07	694	0.04	58.23	2.92	693	0.00
I_{Top} - R	0.11	0.01	692	0.99	58.43	2.74	692	0.01
E_{Bottom} - R	215.67	15.60	672	0.00	128.70	5.99	682	0.00
I_{Bottom} - R	195.83	12.68	680	0.00	139.28	6.01	684	0.00
	IA5							
E_{Top} - I_{Top}	29.06	1.49	690	0.14				
E_{Bottom} - I_{Bottom}	49.35	1.90	664	0.06				
E_{Top} - R	119.94	6.08	689	0.00				
I_{Top} - R	90.88	4.47	697	0.00				
E_{Bottom} - R	296.06	13.18	675	0.00				
I_{Bottom} - R	246.71	10.29	685	0.00				

Table 8

DIC values for each model and each participant. Lowest DIC model is bolded and italicized.

Condition	Participant	Parallel ST	Serial ST	Coactive	Mixed S-P	Free Drift
Upright Aligned	UA1	4963.1	4955.4	4925.9	***4903.3***	4920.8
	UA2	2912.9	2728.7	2756.9	***2622.7***	2734.4
	UA3	1284	1198.4	1096.5	1089	***1039***
	UA4	***4357.6***	4512.2	4391.7	4375.7	4374.8
Upright Misaligned	UM1	3281.1	3318.5	3238.3	***3154.6***	3238.4
	UM2	2891.9	2823.3	2781.6	***2720.4***	2778.3
	UM3	3420.3	3522	3450.7	***3282***	3447.3
	UM4	1047.2	1023.5	1013.2	***1005.5***	1009.1
Inverted Aligned	IA1	-1258.7	-1436.9	-1457.7	***-1529.6***	-1477.8
	IA2	819.09	897.05	893.62	***721.96***	882.64
	IA3	6216.9	6225.6	6171.1	***5950.8***	6189.8
	IA4	-2393.6	-2268.1	-2327.7	***-2498.5***	-2317.6
	IA5	809.11	735.23	666.15	638.08	***630.88***
Inverted Misaligned	IM1	-1638.9	-1744.5	-1807.3	***-1913.8***	-1802.3
	IM2	-1801.2	-1395.9	-1493.6	***-1858.9***	-1476.9
	IM3	2201.3	2353.5	2247.9	***2101.4***	2268.1
	IM4	318.46	-68.554	-78.667	***-350.97***	-58.265

Note: ST = Self-terminating, SP = Serial-Parallel.

Table 9

Summary of Nonparametric and Parametric Results

Participant	SIC	MIC	Interior vs Exterior	Nonparametric Result	Best Fitting Parametric Model
UA1	completely negative	Non-significant	Non-significant	Parallel ST/Serial ST	Mixed S-P
UA2	mostly negative, some late positivity	Non-significant	Interior > Exterior	Serial ST/Parallel ST	Mixed S-P
UA3	mostly negative, some late positivity	Negative	Interior < Exterior	Parallel ST/Coactive	Free Drift
UA4	S-shaped, negative then positive	Non-significant	I < E Top, I > E Bot	Serial ST/Coactive	Parallel ST
UM1	S-shaped, more negative than positive	Negative	Non-significant	Parallel ST	Mixed S-P
UM2	S-shaped, equal negative and positive	Non-significant	Non-significant	Serial ST	Mixed S-P
UM3	completely negative	Negative	Non-significant	Parallel ST	Mixed S-P
UM4	completely negative	Negative	Interior < Exterior	Parallel ST/Coactive	Mixed S-P
IA1	S-shaped, more positive than negative	Non-significant	Interior < Exterior	Serial ST/Coactive	Mixed S-P
IA2	completely negative	Non-significant	Non-significant	Parallel ST/Serial ST	Mixed S-P
IA3	mostly positive	Non-significant	Non-significant	Coactive/Serial ST/Parallel ST	Mixed S-P
IA4	mostly positive	Non-significant	Interior > Exterior	Coactive/Serial ST	Mixed S-P
IA5	S-shaped, mostly positive	Positive	Interior < Exterior	Coactive	Free Drift
IM1	mostly positive	Non-significant	Non-significant	Coactive/Serial ST/Parallel ST	Mixed S-P
IM2	S-shaped, mostly positive	Non-significant	Interior < Exterior	Coactive/Serial ST	Mixed S-P
IM3	S-shaped, mostly positive	Non-significant	Interior > Exterior	Coactive/Serial ST	Mixed S-P
IM4	S-shaped, mostly positive	Non-significant	Non-significant	Coactive/Serial ST/Parallel ST	Mixed S-P

Note: The Nonparametric Result is found by interpreting the SIC, MIC, and Interior vs Exterior results in that order. Where multiple interpretations are possible, all are listed.

Figure Captions

Figure 1. From left to right (top), schematic of the stimulus spaces in Fifić & Townsend (2010)'s OR and AND conditions, and the stimulus space used in the current paper. In the centre, we show the general combination of faces used in the upright aligned face category space for the present experiment; the actual face images were drawn from Kayser (1987) and are available from the authors on request. Dimension X is created by morphing top halves Face 3 and 4. Dimension Y is created by morphing bottom halves Face 1 and 2. We also show an example of the layout from each of the four conditions (bottom).

Figure 2. Left panels: Illustration of the three main patterns of mean RTs for the target category with their corresponding MIC values. L denotes low discriminability on a dimension; H denotes high discriminability on a dimension. Right panels: Illustration of survivor interaction contrasts (SICs) associated with the different processing architectures.

Figure 3. Schematic illustration of contrast category mean RT patterns adapted from Fifić, Little & Nosofsky (2010).

Figure 4. Observed target category SICs (red line) for upright aligned (top row) and upright misaligned (bottom row) participants. 95% bootstrapped CI's (blue lines) are also shown. The small inset figures show the target category mean RTs. The two left hand points have low discriminability on the top half, and the two right-hand points have high discriminability on the top half. The solid line (with black marker) has low discriminability on the bottom half; the dotted line (with white marker) has high discriminability on the bottom half.

Error bars represent ± 1 SE. Some of the standard error bars for the mean RTs are too small to be seen.

Figure 5. Observed contrast category mean RTs for the upright aligned (top row) and upright misaligned (bottom row) participants. Error bars represent ± 1 SE. Some of the standard error bars for the mean RTs are too small to be seen. The label “Top” refers to items which satisfy the category B rule on the Top face half (i.e., the vertical boundary) but vary on the bottom face half values. Correspondingly, the label “Bottom” refers to items which satisfy the category B rule on the bottom face half (i.e., the horizontal boundary) but vary on the top face half values.

Figure 6. Observed target category SICs (red line) for inverted aligned (top row) and inverted misaligned (bottom row) participants. 95% bootstrapped CI’s (blue lines) are also shown. The small inset figures show the target category mean RTs. The two left hand points have low discriminability on the top half, and the two right-hand points have high discriminability on the top half. The solid line (with black marker) has low discriminability on the bottom half; the dotted line (with white marker) has high discriminability on the bottom half. Error bars represent ± 1 SE. Some of the standard error bars for the mean RTs are too small to be seen.

Figure 7. Observed contrast category mean RTs for the inverted aligned (top row) and inverted misaligned (bottom row) participants. Error bars represent ± 1 SE. Some of the standard error bars for the mean RTs are too small to be seen. The label “Top” refers to items which satisfy the category B rule on the Top face half (i.e., the vertical boundary) but vary on the bottom face half values. Correspondingly, the label “Bottom” refers to items which satisfy the

category B rule on the bottom face half (i.e., the horizontal boundary) but vary on the top face half values.

Figure 8. Panel A: GRT representation assumed by the logical rule models. The category space is represented by nine bivariate normal distributions. The marginal distributions on each dimension are shaded. Panel B: LBA accumulator pairs for the top and bottom face halves. Panel C: Example of serial, parallel, and coactive processing architectures as applied to a target category stimulus.

Figure 9. Mean posterior predictions of the mixed serial-parallel model for the 25, 50, 75% percentiles for the correct and error RTs. Each point represents a single item from a single subject from the Upright Aligned (x's), Upright Misaligned (o's), Inverted Aligned (Δ's), and Inverted Misaligned (*'s).

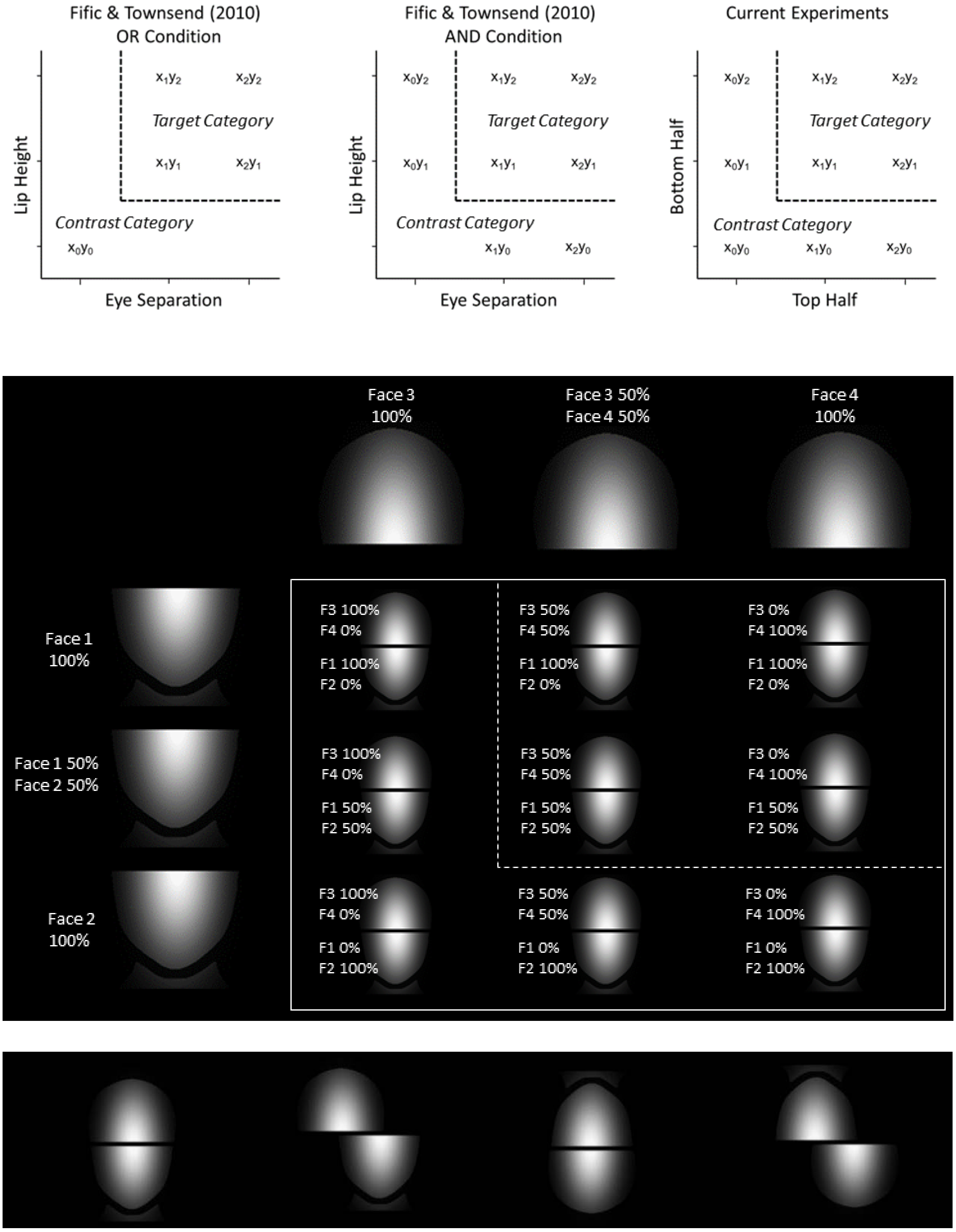
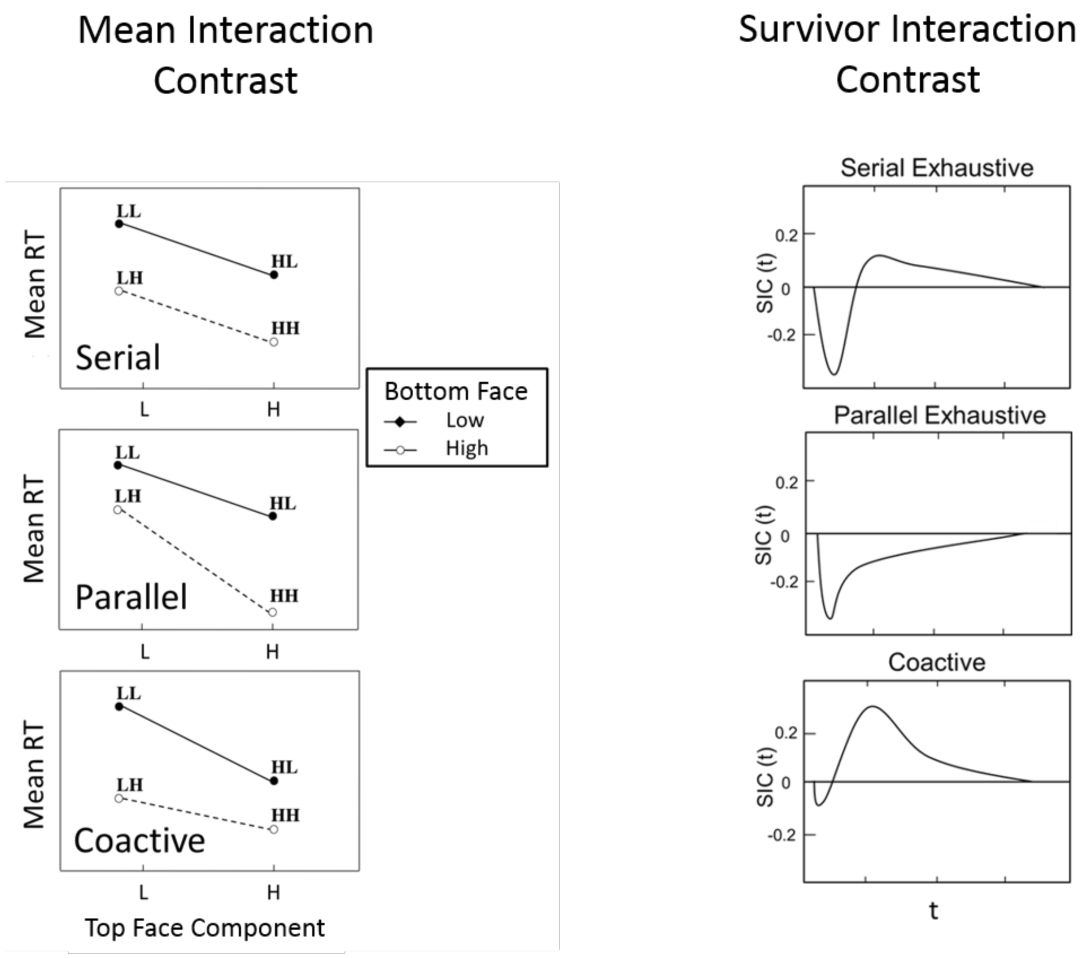

Fific & Townsend (2010)
OR Condition
Lip Height
Eye Separation
x_1y_2 x_2y_2
Target Category
x_1y_1 x_2y_1
Contrast Category
x_0y_0
Fific & Townsend (2010)
AND Condition
Lip Height
Eye Separation
x_0y_2 x_1y_2 x_2y_2
Target Category
x_0y_1 x_1y_1 x_2y_1
Contrast Category
x_1y_0 x_2y_0
Current Experiments
Bottom Half
Top Half
x_0y_2 x_1y_2 x_2y_2
Target Category
x_0y_1 x_1y_1 x_2y_1
Contrast Category
x_0y_0 x_1y_0 x_2y_0
Face 3
100%
Face 3 50%
Face 4 50%
Face 4
100%
Face 1
100%
Face 1 50%
Face 2 50%
Face 2
100%
F3 100%
F4 0%
F1 100%
F2 0%
F3 50%
F4 50%
F1 100%
F2 0%
F3 0%
F4 100%
F1 100%
F2 0%
F3 100%
F4 0%
F1 50%
F2 50%
F3 50%
F4 50%
F1 50%
F2 50%
F3 0%
F4 100%
F1 50%
F2 50%
F3 100%
F4 0%
F1 0%
F2 100%
F3 50%
F4 50%
F1 0%
F2 100%
F3 0%
F4 100%
F1 0%
F2 100%

Figure 1.



Mean Interaction Contrast
Survivor Interaction Contrast
Mean RT
LL
HL
LH
HH
Serial
L
H
Bottom Face
Low
High
Parallel
Coactive
Top Face Component
Serial Exhaustive
Parallel Exhaustive
SIC (t)
0.2
0
-0.2
t

Figure 2.

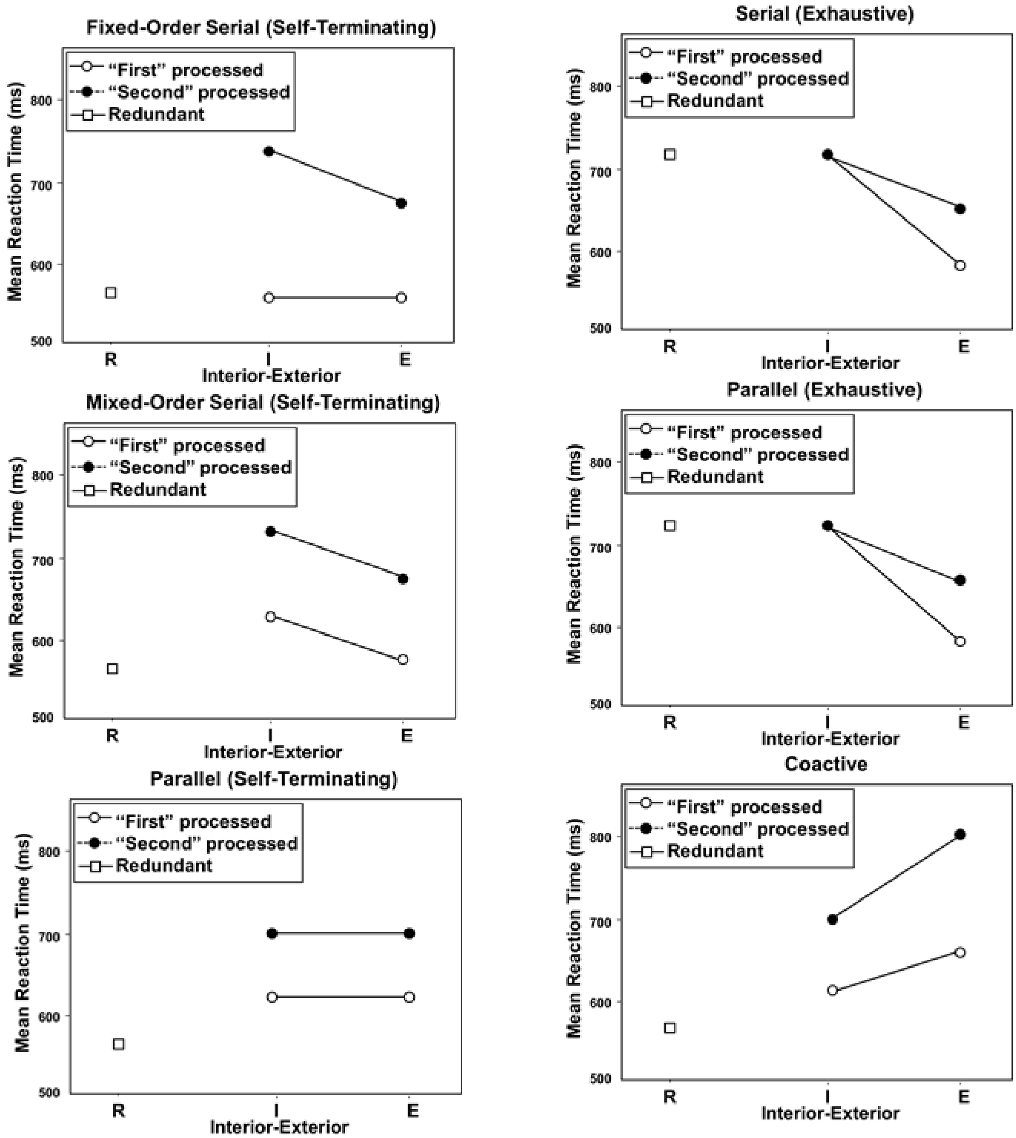

Fixed-Order Serial (Self-Terminating)
Mixed-Order Serial (Self-Terminating)
Parallel (Self-Terminating)
Serial (Exhaustive)
Parallel (Exhaustive)
Coactive
"First" processed
"Second" processed
Redundant
Mean Reaction Time (ms)
Interior-Exterior
R
I
E
500
600
700
800

Figure 3.

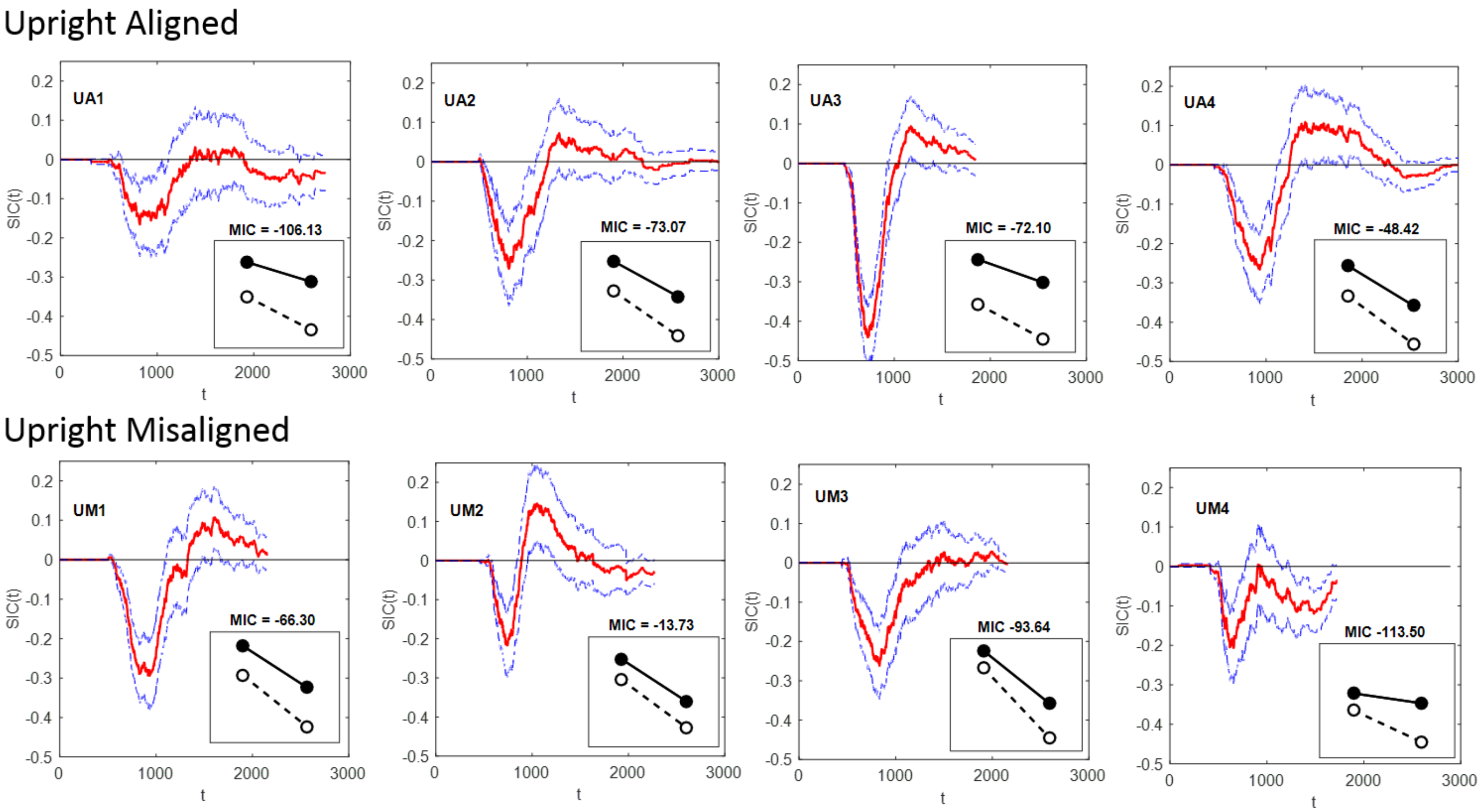

Upright Aligned
UA1
MIC = -106.13
UA2
MIC = -73.07
UA3
MIC = -72.10
UA4
MIC = -48.42
Upright Misaligned
UM1
MIC = -66.30
UM2
MIC = -13.73
UM3
MIC -93.64
UM4
MIC -113.50
SIC(t)
t

Figure 4.

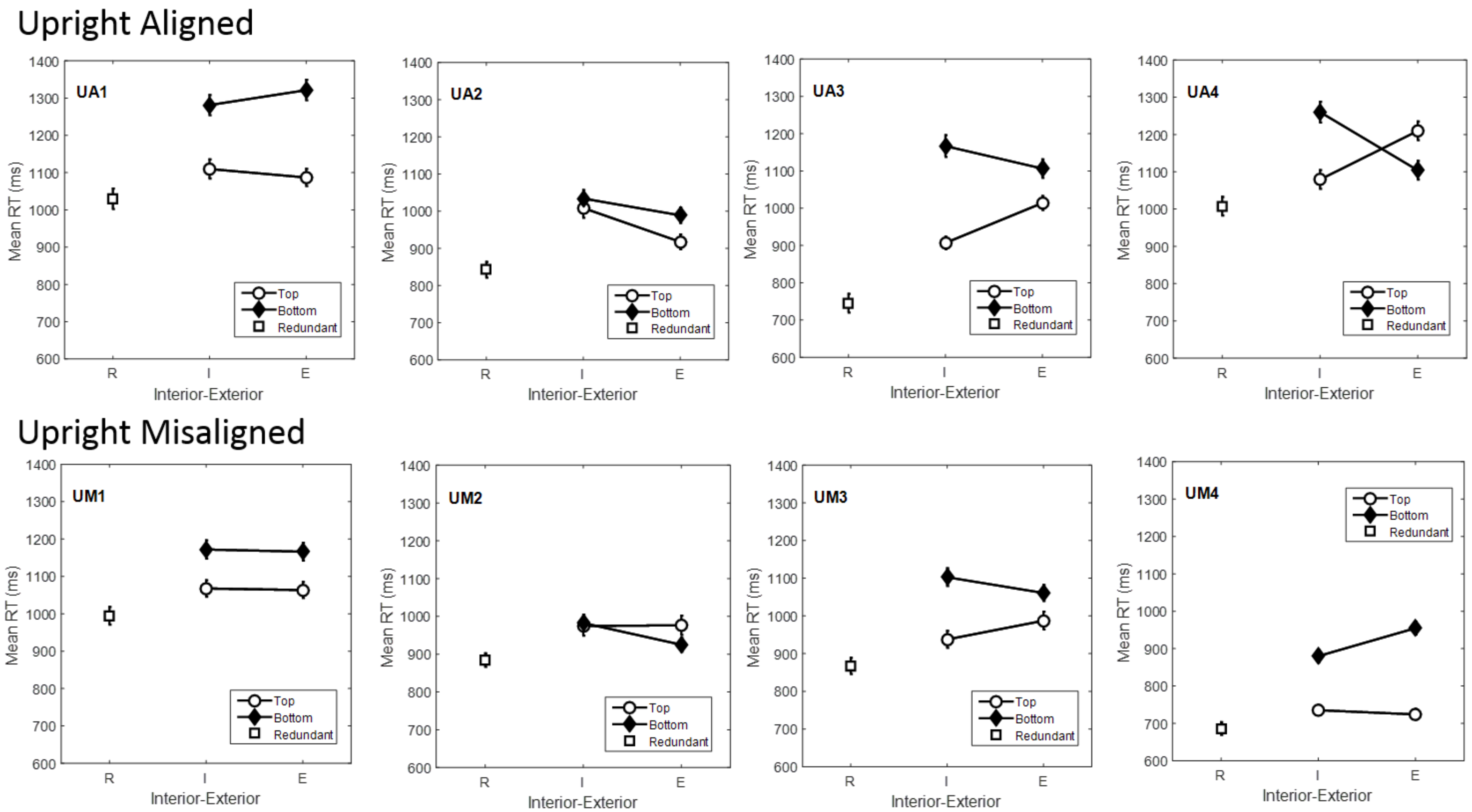
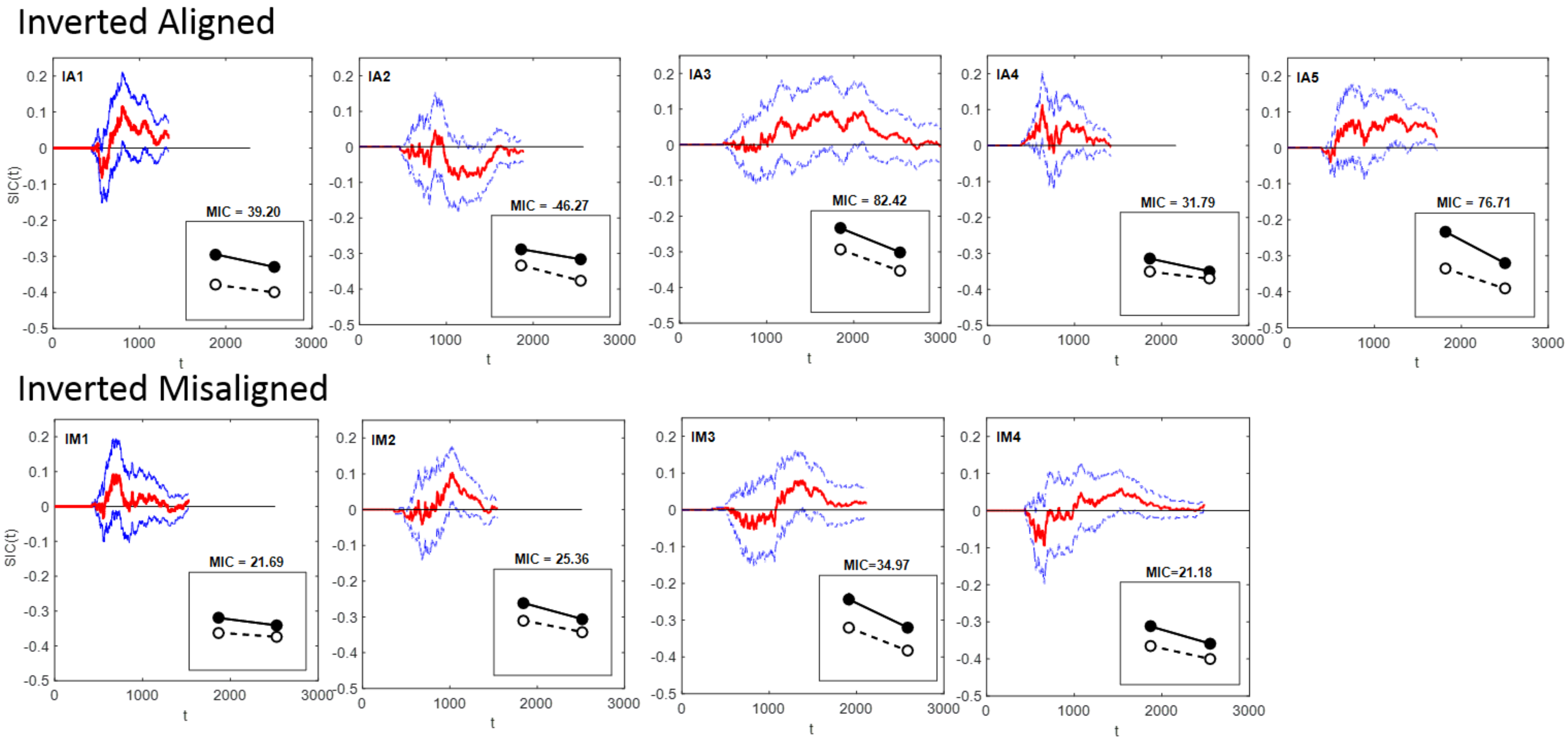

Upright Aligned
UA1
UA2
UA3
UA4
Upright Misaligned
UM1
UM2
UM3
UM4
Mean RT (ms)
Interior-Exterior
R
I
E
Top
Bottom
Redundant
600
700
800
900
1000
1100
1200
1300
1400

Figure 5.



Inverted Aligned
IA1
MIC = 39.20
IA2
MIC = -46.27
IA3
MIC = 82.42
IA4
MIC = 31.79
IA5
MIC = 76.71
Inverted Misaligned
IM1
MIC = 21.69
IM2
MIC = 25.36
IM3
MIC=34.97
IM4
MIC=21.18
SIC(t)
t

Figure 6.

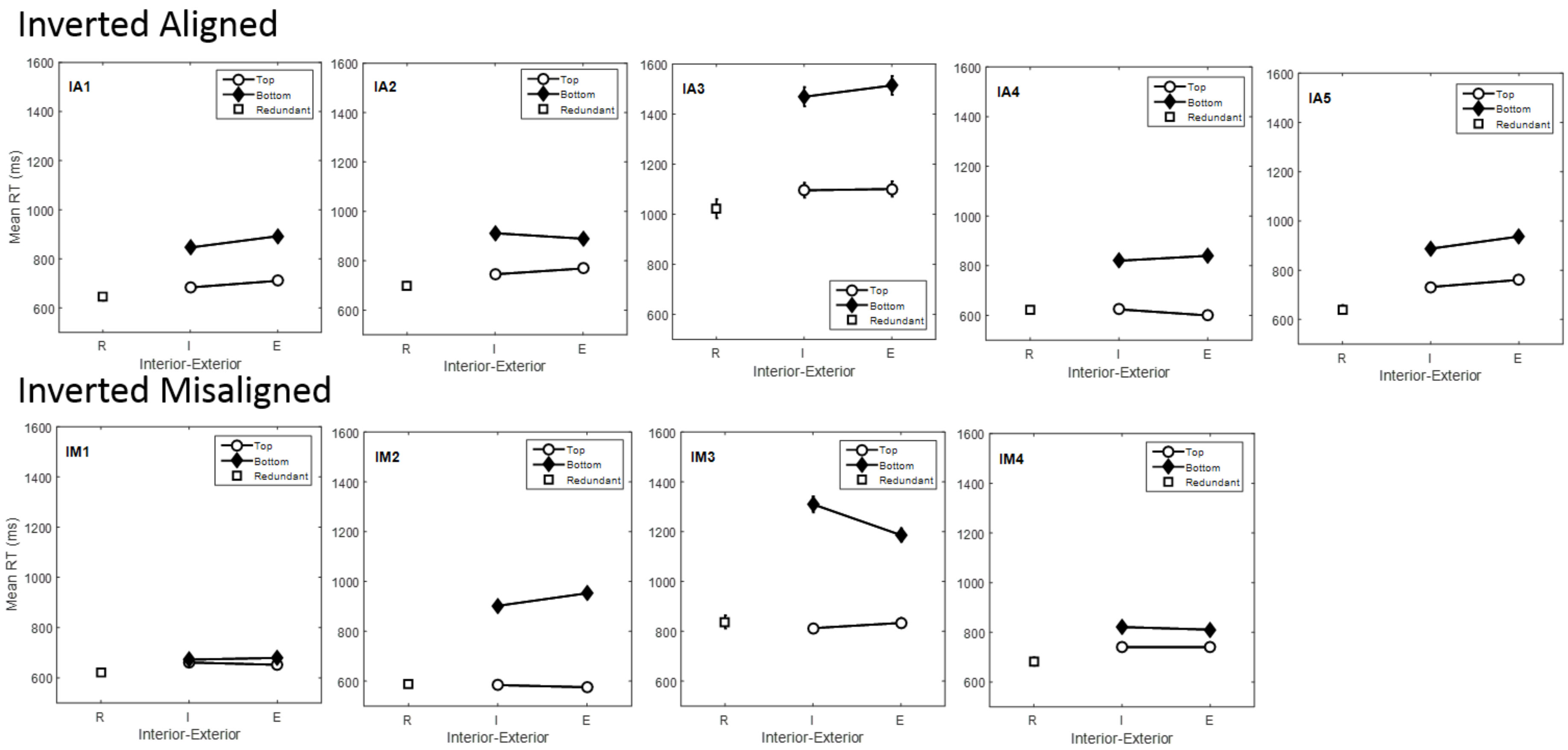

Inverted Aligned
IA1
IA2
IA3
IA4
IA5
Top
Bottom
Redundant
Mean RT (ms)
1600
1400
1200
1000
800
600
R
I
E
Interior-Exterior
Inverted Misaligned
IM1
IM2
IM3
IM4

Figure 7.

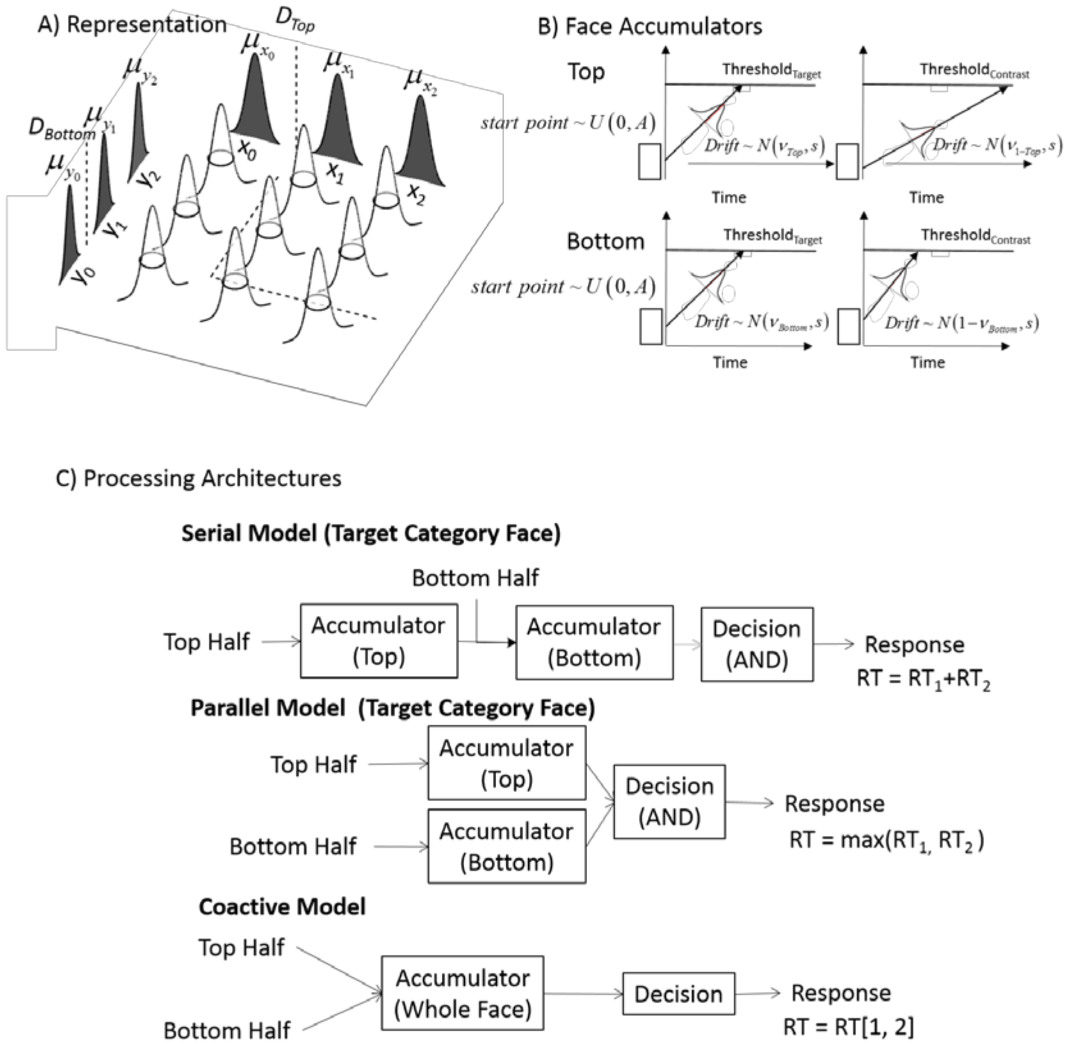

A) Representation
B) Face Accumulators
Top
Bottom
Threshold_Target
Threshold_Contrast
start point ~ U(0, A)
Drift ~ N(v_Top, s)
Drift ~ N(v_1-Top, s)
Drift ~ N(v_Bottom, s)
Drift ~ N(1-v_Bottom, s)
Time
C) Processing Architectures
Serial Model (Target Category Face)
Bottom Half
Top Half
Accumulator (Top)
Accumulator (Bottom)
Decision (AND)
Response
RT = RT1+RT2
Parallel Model (Target Category Face)
Top Half
Bottom Half
Accumulator (Top)
Accumulator (Bottom)
Decision (AND)
Response
RT = max(RT1, RT2)
Coactive Model
Top Half
Bottom Half
Accumulator (Whole Face)
Decision
Response
RT = RT[1, 2]

Figure 8.

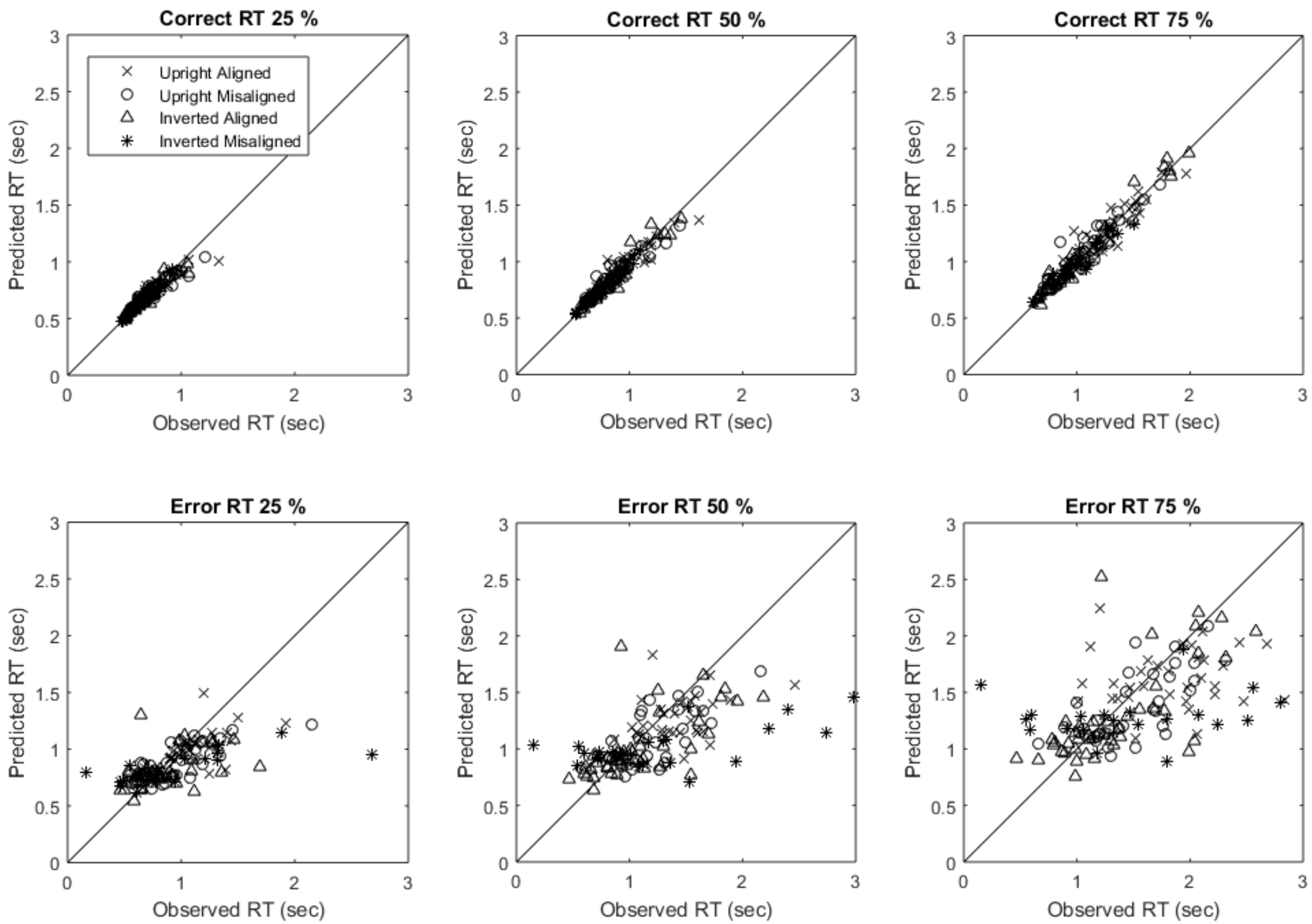

Correct RT 25 %
Correct RT 50 %
Correct RT 75 %
Error RT 25 %
Error RT 50 %
Error RT 75 %
Upright Aligned
Upright Misaligned
Inverted Aligned
Inverted Misaligned
Predicted RT (sec)
Observed RT (sec)
0
0.5
1
1.5
2
2.5
3
0
1
2
3

Figure 9.